Enhanced Radio Propagation Modelling for Future Wireless Networks

Lawrence Gregory Sayer

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Doctor of Philosophy in the Faculty of Engineering.

Entrust your works to the Lord, and your plans will succeed.

Proverbs 16:5

Dedication

To Siobhán, Stanley, Esther, *Mater* and *Pater*.

Declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author. This work was funded by the EPSRC and the University of Bristol.

Signature _____ Date _____

Copyright Notice

Abstract

This thesis relates to the subject of propagation modelling for cellular radio communications. Since current generations of cellular radio are forecast to be incapable of delivering the data traffic expected by consumers, new 5G radio specifications have been put forward which have stringent latency, data rate, density and mobility requirements. To meet these requirements new technologies have been proposed such as mmWave and MIMO. To obtain a sufficient understanding of wireless radio channel behaviour to design these new systems well, channel models are required that handle their idiosyncrasies. This thesis discusses propagation modelling theory in Chapter 2, and shows how a ray tracing model can be used to design communication systems in Chapter 3. In Chapter 4, a new ray launching tool, designed to be an experimental platform for investigating models of propagation phenomena, is documented. In Chapter 5, diffuse scattering and diffraction models are implemented and evaluated against measured data. Chapter 6 describes techniques for the optimisation of ray tracing tools, and a new ray tracing tool is implemented as part of this work. Chapter 7 is devoted to the radio channel in situations in which there is mobility.

Contents

List of Figures

List of Tables

Chapter 1

Motivation and Contributions

1.1 Motivation for Next Generation Cellular Radio – 5G

The fifth generation of cellular radio, 5G, will deliver a wide range of services to consumers. Different radio technologies will be used to meet the various Quality of Service (QoS) requirements for each these new services [13]. Demand for wireless broadband data is driven by digital applications such as video on-demand, instant-messaging and social media. As computing devices become more sophisticated, the applications that run on them continue to demand yet more data. Furthermore, new use-cases for wireless data transfer have been proposed, including machine to machine communications and ultra-reliable low-latency vehicular communication.

There has been an evolution of wireless cellular networks. The 'G' for 'generation' nomenclature has been applied to group the main developments. 1G described analogue networks designed for voice. 2G added text messaging and low data rate data. 3G enabled internet access with higher data rates and, most recently, 4G has continued to provide higher and higher data rates to meet increasing consumer demand for data. However, with a predicted 1000 fold increase in demand for data, the current networks are forecast to be overstretched by 2020 [16].

Type of performance requirement	Usage scenario	KPI (minimum requirement)
Peak data rate	eMBB	Downlink peak data rate: 20 Gbit/s. Uplink peak data rate: 10 Gbit/s.
Peak spectral efficiency	eMBB	Downlink peak spectral efficiency: 30 bit/s/Hz. Uplink peak spectral efficiency: 15 bit/s/Hz.
User experienced data rate	eMBB	Downlink user experienced data rate: 100 Mbit/s. Uplink user experienced data rate: 50 Mbit/s.
Area traffic capacity	eMBB	Downlink: 10 Mbit/s/m2 in the Indoor Hotspot – eMBB test environment.
Connection density	mMTC	1 000 000 devices per km2
Mobility	eMBB	Indoor hotspot: Stationary, Pedestrian* Dense Urban: Stationary, Pedestrian, Vehicular (up to 30 km/h)* Rural: Pedestrian, Vehicular, High speed vehicular (up to 500 km/h)*
Bandwidth	IMT-2020	at least 100 MHz up to 1 GHz for operation in higher frequency bands (e.g. above 6 GHz).

Table 1.1: 5G performance requirements

A subset of requirements for 5G technology are presented in Table 1.1 [1]. In the table, the acronym eMBB stands for enhanced mobile broadband and mMTC stands for massive machine type communication. A third key area of focus for 5G is uRLLC (ultra reliable low latency communication). Peak data rates of 20 Gb/s have been specified - there will also be a higher density of devices, with up to 1 000 000 devices per square kilometre and a sharp decrease in allowable latency, down to 1 ms for some applications [21]. To appropriately design 5G networks and devices to meet such re-

quirements, it is necessary to understand the wireless radio channel. Insight into the channel can not be gained from tangible experience, rather engineers depend upon, either, tools that simulate the channel, or real world measurements of it. Measurements can be difficult, limited in scope and costly, which makes modelling an attractive proposition. There are already a wide range of models available. However, for 5G, new radio technologies will be used and therefore there will be a shift in the importance of different propagation phenomena, and the data that need to be extracted from models will be different as well. Form a propagation modelling perspective, this presents a challenge and existing models are not anticipated to be suitable.

There are many potential applications of good propagation models. These include; planning secure or domain limited networks [22]; simulating 5G technologies in indoor environments [17]; simulating mmWave and Multiple Input and Multiple Output (MIMO) links [2][5]; simulation of highly dynamic IoT networks [9]; and generating training data for or assisting localisation algorithms [20].

1.1.1 Millimetre Wavelength (mmWave)

The spectrum bands currently used for cellular radio are fragmented, expensive and congested. Using millimetre wavelengths to communicate allows unlicensed spectrum to be used. This means wider continuous bandwidths are available, up to 1 GHz in some cases [21]. This will facilitate higher data rates and lower latency wireless links [16]. Small cells can be used at extremely high data rates with low interference between cells. The minimisation of components also means more antennas may be fit on devices, which can then reap the benefits provided by exploiting spatial diversity [15]. The use of high gain beamforming techniques makes mmWave links more practical because they offset the relatively high loss inherent to the frequencies involved. In fact, the short wavelength of mmWave links makes it especially easy to apply various spacial processing techniques [16]. The links may also provide back-haul services, due to their very high data rate capability, allowing cable or fibre networks to be entirely replaced where this is an advantage [16]. The problems to be overcome include the difficulty of manufacture of very small devices with tight tolerances, and the significant attenuation of these frequencies when penetrating building materials and the atmosphere [15].

Since these frequencies have rarely been used in the past for cellular radio, mmWave propagation has not been investigated in depth. There are several propagation phenomena that are more or less pronounced at mmWave frequencies when compared to traditional sub-6 GHz propagation. Diffraction is less efficient. Diffuse scattering occurs from surfaces that are rough relative to the shorter wavelengths involved. Fig 1.1 shows both the changing shape of the diffuse scattering radiation pattern with increasing roughness and the reason for this change. When bumps and troughs on a surface are much smaller than a wavelength, plane wave radiation impinging on the surface is reflected, parallel rays arriving at the surface have their phase relationship maintained.

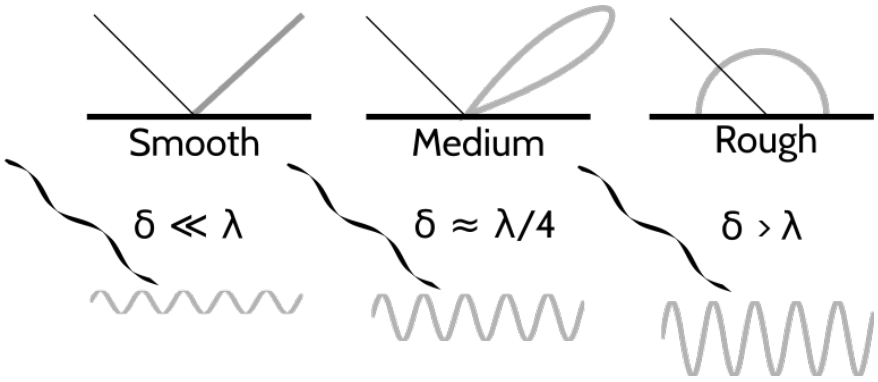


Figure 1.1: Diffuse scattering of incident radiation is caused by surface roughness.

As the surface becomes rougher, the rays that must propagate down to the troughs travel further than those that reflect of the peaks. This means that a phase offset between once co-phased parts of a wavefront is introduced. These phase offsets cause radiation to be scattered in non-specular directions. The effect becomes more and more extreme as the surface continues to increase in roughness. Atmospheric attenuation occurs due to humidity or precipitation. Path losses are high because miniaturised antennas have small effective areas [18]. This means that models calibrated to perform well for sub-6 GHz networks are unlikely to work 'as is' for planning mmWave networks.

1.1.2 MIMO Techniques

To meet the capacity requirements for 5G cellular, MIMO is considered to be a key enabling technology [10]. MIMO exploits the multipath characteristics of the channel by using multiple transmission and receive antennas. Arrays may apply precoding [8] to achieve spatial multiplexing in which a signal is split it to multiple lower rate streams. If these streams have different enough spatial signatures, and the Channel State Information (CSI) is known at the receiver, then these streams can be separated into almost parallel channels [3]. This concept is illustrated in Fig 1.2. Fig 1.2 shows a situation in which an array has a line of sight to three users. It uses precoding to estimate the channel and then is able to steer a beam at each user using the same time and frequency resources. In a scenario with a richer multipath structure, the 'beams' would not have an obvious directional nature but nonetheless the desired signal would constructively combine at each users location.

Since in MIMO schemes many proximate array elements are used in unison to transmit a signal in the desired way, some interesting modelling questions are raised. Since MIMO relies upon having different spatial signatures at each array element, the variations in amplitude and phase of a received signal across an array is of interest. This causes some potential problems. Modelling often injects errors due to the artefacts and idiosyncrasies of the modelling process. An example would be a path length error - and hence phase error - due to the discretisation of the environment in a model. To retrieve reliable statistics about actual spatial signatures may be difficult in such a

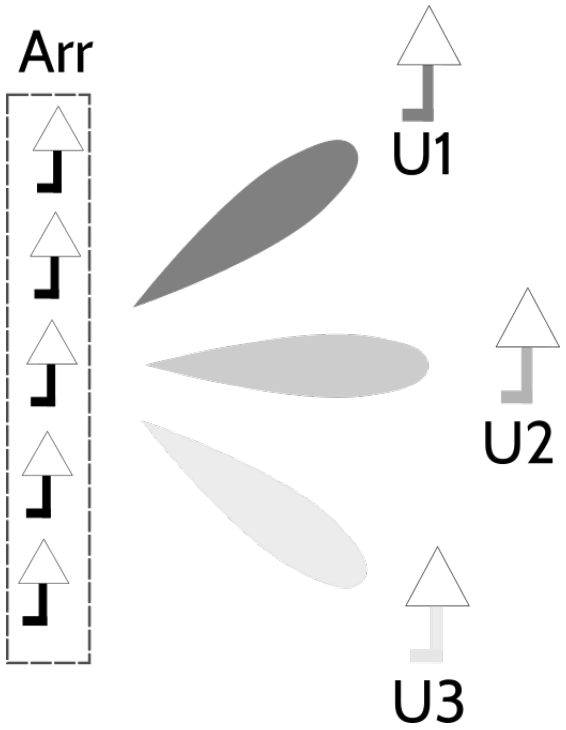


Figure 1.2: Beamforming using an antenna array.

case. Furthermore, there are many elements on an array and due to the large number of elements at both ends of the link, many point to point links may have to be simulated. This is costly in terms of execution time. There has been limited research in this domain.

1.1.3 Vehicular Channels

The use of vehicle to infrastructure or vehicle to vehicle links (V2X) promises to allow cars to become driverless, improving comfort and safety for road users [4]. Both industrial applications, such as convoys of freight traffic, and consumer grade applications, such as in car entertainment are proposed. V2X networks are heavily laden with strict QoS requirements to ensure safety. To allow vehicles to navigate roads and manoeuvre around obstacles such as other cars, pedestrians and hazards, very low latency, ultra reliable, high data rate links are required [12]. Several challenges are faced by designers of networks containing connected autonomous vehicles (CAVs). Network topology can change very fast, and radios are subject to intermittent blockages from other vehicles, buildings and infrastructure [11]. To model such networks, and especially to understand the channel dynamics with moving objects in a quasi-optical network, new models are required.

1.1.4 Internet of Things (IoT)

By the year 2021, it is predicted that there will be around 28 billion connected devices [7]. A large proportion of such devices will be industrial devices with an internet connection. Feeding sensor

data back to process monitors, either human or machine, will enable industrial processes to be better understood and therefore be better optimised, saving energy labour and expense and performing tasks to a higher standard [19].

One of the challenges for industrial IoT applications is the wide range of requirements that different devices will have [23]. The data output of a robotic-arm-mounted camera, streaming video in real time to an image processing application, has very different requirements to a low-cost low-power agricultural sensor that reports humidity levels once every day. Furthermore, IoT networks will possibly exist in extreme or dangerous environments, and are likely to be designed to be low cost - which is one of the prerequisites of their ubiquity. From a propagation perspective, understanding the effects of low manufacturing quality, extreme conditions and low power is essential. A key to the social and economic benefits promised by the IoT, is well informed device design based on sound modelling tools [14][6].

1.2 Contributions

Over the course of the research presented in this thesis, several contributions to the field have been made. A summary is presented here.

1. An iterative antenna design and analysis method that makes use of ray tracing simulations. The method is the subject of Chapter 2.

2. The design and implementation of a high performance vector-based 3D ray launching tool, including a method for finding initial ray directions. Crucially, this allowed new and experimental propagation models to be tested. The design is described in Chapter 4. This tool was used in Chapters 5 and 7 to obtain results relating to 5G specific networks.

3. A new method for finding multiply diffracted ray paths in 3D space for a ray launcher. This method is presented in Chapter 5.

4. Implementation, analysis and validation of scattering and diffraction models for millimetre wavelength propagation. This involved two sophisticated and highly controlled measurement campaigns. This work is described in Chapter 5. Modelling diffuse scattering included addition of a statistically varying component to a Kirchhoff scattering model to better recreate small scale fading effects.

5. Development of a tracing tool specifically designed for simulation of MIMO networks. The tool was designed to accurately predict phase relationships across both space and time, facilitating the use of channel data to benchmark spatial diversity or spatial multiplexing schemes. The tool is presented in Chapter 6.

6. A robust method for finding the intra-mesh visibility relations for a 3D model of a city, greatly increasing the speed of ray tracing simulations. This is presented in Chapter 6. Some intelligent, but ultimately unsuccessful, attempts to further optimise ray tracing for antenna arrays were also made. These methods and analysis of their failure is also presented.

7. Simulation of networks containing mobility, including Doppler shift calculation, to the ray launching tool. This enables the Doppler spectrum to be extracted from ray launching simulations. The method for calculation of Doppler shifts, and their use to analyse OTFS networks is presented in Chapter 7.

1.3 Publications List

1. Sayer, L, Freire, AL, Mellios, E & Nix, A, 2018, 'A Kirchhoff Scattering Model for Millimetre Wavelength Wireless Links'. in: *The 12th European Conference on Antennas and Propagation*. (Eucap 2018).

2. Freire, AL, Pelham, T, Kong, D, Sayer, L, Sgardoni, V, Tila, F, Mellios, E, Beach, M, Nix, A & Steinbock, G, 2018, 'Polarimetric Diffuse Scattering Channel Measurements at 26 GHz and 60 GHz'. in: *2018 IEEE 29th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications* (PIMRC)-Workshop WS-07 on "Millimiter Waves Communications": IEEE PIMRC 2018 Workshop WS-07 on "Millimiter Waves Communications".

3. Dumanli, S, Sayer, L, Mellios, E, Fafoutis, X, Hilton, G & Craddock, I, 2017, 'Off-body antenna wireless performance evaluation in a residential environment'. *IEEE Transactions on Antennas and Propagation*, vol 65., pp. 6076-6084

4. Wiffen, F, Sayer, L, Bocus, Z, Doufexi, A & Nix, A, 2018, 'Comparison of OTFS and OFDM in Ray Launched sub-6 GHz and mmWave Line-of-Sight Mobility Channels'. in: *2018 IEEE 29th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications* (PIMRC 2018). Institute of Electrical and Electronics Engineers (IEEE), pp. 73-79

5. Sayer, L, Mellios, E, & Nix, A, 2019, 'On the Feasibility of NLOS Vehicular Communication via Rough Surfaces at 28 GHz'. (Accepted) *IEEE Global Communications Conference* (Globecom 2019). Institute of Electrical and Electronics Engineers (IEEE).

6. Sayer, L, Nix, A, & Beach, M, 2019 'Method for calculation of intra-mesh visibility matrix for enhanced 3D ray tracing simulations', *IET Microwaves, Antennas &; Propagation* (Ready for Submission) IET Digital Library, `https://digital-library.theiet.org`.

7. Sayer, L, Freire, AL, 2019, 'mmWave Propagation Around Rough and Smooth Conducting Wedges' (Submitted) in: *The 2020 IEEE 91st Vehicular Technology Conference*. (VTC Spring 2020).

Bibliography

[1] 5G-ppp. D2.6 - final report on programme progress and KPIs.

[2] Zahid Aslam, Yoann Corre, Emil Björnson, and Erik G. Larsson. Performance of a dense urban massive MIMO network from a simulated ray-based channel. *EURASIP Journal on Wireless Communications and Networking*, 2019, 12 2019.

[3] A. W. C. Lim and V. K. N. Lau. On the fundamental tradeoff of spatial diversity and spatial multiplexing of MIMO links with imperfect CSIT. In *2006 IEEE International Symposium on Information Theory*, pages 2704–2708, July 2006.

[4] M. Daily, S. Medasani, R. Behringer, and M. Trivedi. Self-driving cars. *Computer*, 50(12):18–23, December 2017.

[5] V. Degli-Esposti, F. Fuschini, E. M. Vitucci, M. Barbiroli, M. Zoli, L. Tian, X. Yin, D. A. Dupleich, R. Müller, C. Schneider, and R. S. Thomä. Ray-tracing-based mm-Wave beamforming assessment. *IEEE Access*, 2:1314–1325, 2014.

[6] S. Dumanli, L. Sayer, E. Mellios, X. Fafoutis, G. S. Hilton, and I. J. Craddock. Off-body antenna wireless performance evaluation in a residential environment. *IEEE Transactions on Antennas and Propagation*, 65(11):6076–6084, Nov 2017.

[7] Ericsson. *Cellular networks for massive IoT. [Online].* Januaray 2016.

[8] N. Fatema, G. Hua, Y. Xiang, D. Peng, and I. Natgunanathan. Massive MIMO linear precoding: A survey. *IEEE Systems Journal*, 12(4):3920–3931, Dec 2018.

[9] S. Hosseinzadeh, H. Larijani, K. Curtis, A. Wixted, and A. Amini. Empirical propagation performance evaluation of LoRa for indoor environment. In *2017 IEEE 15th International Conference on Industrial Informatics (INDIN)*, pages 26–31, July 2017.

[10] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta. Massive MIMO for next generation wireless systems. *IEEE Communications Magazine*, 52(2):186–195, February 2014.

[11] N. Lu, N. Cheng, N. Zhang, X. Shen, and J. W. Mark. Connected vehicles: Solutions and challenges. *IEEE Internet of Things Journal*, 1(4):289–299, Aug 2014.

[12] I. Mavromatis, A. Tassi, R. J. Piechocki, and A. Nix. Efficient V2V communication scheme for 5G mmWave hyper-connected CAVs. In *2018 IEEE International Conference on Communications Workshops (ICC Workshops)*, pages 1–6, May 2018.

[13] Amy Nordrum and Kristen Clark. Everything you need to know about 5G.

[14] T. O. Olasupo. Propagation modeling of IoT devices for deployment in multi-level hilly urban environments. In *2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, pages 346–352, Nov 2018.

[15] R. T. Prabu, M. Benisha, V. T. Bai, and V. Yokesh. Millimeter wave for 5G mobile communication application. In *2016 2nd International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB)*, pages 236–240, Feb 2016.

[16] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez. Millimeter wave mobile communications for 5G cellular: It will work! *IEEE Access*, 1:335–349, 2013.

[17] Remcom. *5G mmWave Channel Modeling with Diffuse Scattering in an Office Environment. ebook*. (n.d.).

[18] M. Shafi, J. Zhang, H. Tataria, A. F. Molisch, S. Sun, T. S. Rappaport, F. Tufvesson, S. Wu, and K. Kitao. Microwave vs. millimeter-wave propagation channels: Key differences and impact on 5G cellular systems. *IEEE Communications Magazine*, 56(12):14–20, December 2018.

[19] E. Sisinni, A. Saifullah, S. Han, U. Jennehag, and M. Gidlund. Industrial internet of things: Challenges, opportunities, and directions. *IEEE Transactions on Industrial Informatics*, 14(11):4724–4734, Nov 2018.

[20] A. Tayebi, J. Gomez, F. S. de Adana, and O. Gutierrez. Ray-tracing application to mobile localization in multipath indoor environments. In *2009 International Conference on Electromagnetics in Advanced Applications*, pages 412–415, Sep. 2009.

[21] International Telecommunications Union. Recommendation itu-r m.2083-0, imt vision – framework and overall objectives of the future development of imt for 2020 and beyond. pages 1–21, September 2015.

[22] C. Wang and H. Wang. Physical layer security in millimeter wave cellular networks. *IEEE Transactions on Wireless Communications*, 15(8):5569–5585, Aug 2016.

[23] L. D. Xu, W. He, and S. Li. Internet of things in industries: A survey. *IEEE Transactions on Industrial Informatics*, 10(4):2233–2243, Nov 2014.

Chapter 2

Propagation

2.1 The Wireless Radio Propagation Channel

In a communications system, a data source attempts to convey information to a data sink by transmitting it over a channel. The channel modifies the signal produced by the transmitter in an unpredictable way, making the deciphering of information at the receiver more difficult. The interaction of the channel proper and the antenna exciting the medium is complex. The composite of the antenna and the propagation environment must be considered.

Fig 2.1 shows a wireless communications system with noise in the channel. Additive noise is present due to cosmic radiation, interfering sources, and effects in the receiver itself such as thermal and shot noise. Multiplicative noise is generated by fast and slow fading and path loss acting on the signal between the two antennas - whose radiation patterns also have a multiplicative effect on the received signal. In a wireless radio channel, there is also a degradation of power as emitted radio waves spread through space and move through objects that attenuate them. The primary objective of the wireless system designer is to mitigate as best as possible the effect of these sources

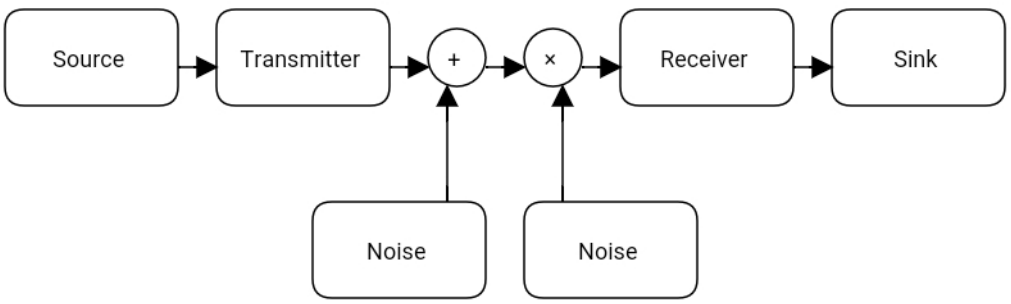


Figure 2.1: A wireless communication system.

of noise, interference and signal degradation on the timely and reliable transfer of information between source and sink. In the wireless context this normally means that it is desired to maximise the power of the wanted signal at the RX.

Various properties of the channel may be of interest to network and device designers. For early broadcast and military applications, losses over long distances were of most importance. Where all users receive the same signal, or where, in special circumstances, a mission critical signal must be transmitted at all costs over long distances, with little regard for coexisting entities, simple power predictions are all that is needed. In cellular radio systems, interference between gracefully coexisting users and cells needed to be understood. Propagation prediction in multipath situations became crucial to network planning and impacted the design of devices and protocols used to transmit voice and data. More recently there has been a drive to use inherently "spatial" communications techniques such as beamforming or MIMO. This is because, where traditionally time and frequency have been the domains over which multiple access has been achieved, these resources are limited. The new emphasis on space is motivated by the need to achieve higher capacity using constrained resources in time and frequency. Fig 2.2 shows how adding a spatial dimension allows the same frequency and time resources to be used for a multiple users, increasing the capacity of a system with constrained time frequency resources.

2.2 Electromagnetic Waves

2.2.1 Introduction

Electromagnetic waves were described by Maxwell's equations in 1865 [7]. Examining Maxwell's formulations of Fraday's law (2.1)

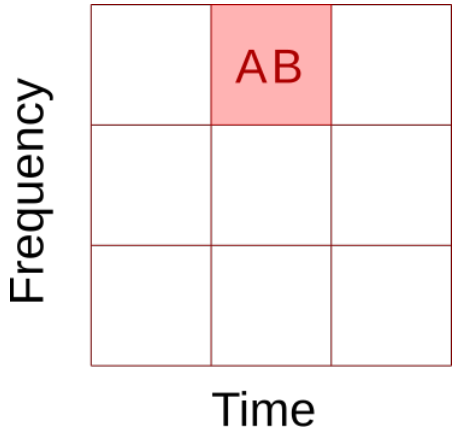
$$\nabla \times \mathrm{E} = -\frac{\delta \mathrm{B}}{\delta t}, \tag{2.1}$$

and Ampere's Law (2.2)

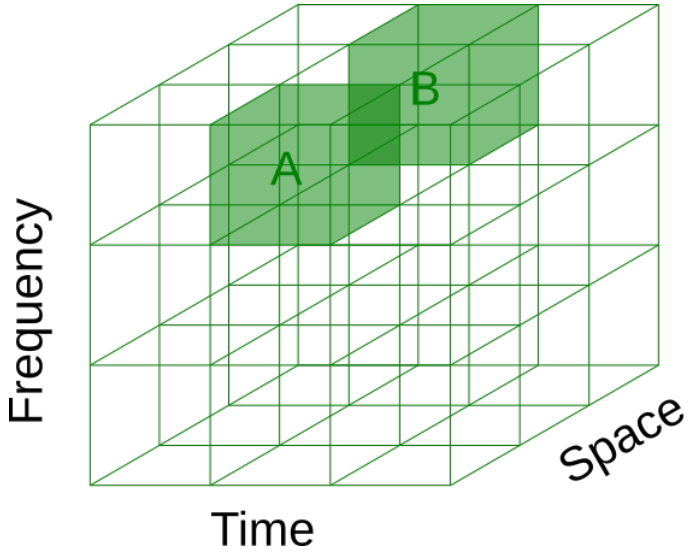
$$\nabla \times \mathrm{H} = \frac{\delta \mathrm{D}}{\delta t} + \mathrm{J}, \tag{2.2}$$

it can be observed that a changing magnetic field generates a changing electric field and a changing electric field generates a changing magnetic field [7]. By this repeated mechanism, energy can propagate through space, allowing information to be communicated by modulating the propagating wave.

Electromagnetic waves have both electric field, E, and magnetic field, H, components which exist in planes perpendicular to one another and the direction of propagation [9]. Propagation speed is dependent on the medium through which the wave propagates. Fig 2.3 shows a propagating electromagnetic wave with perpendicular electric and magnetic field components.



(a) Reusing the same time and frequency resource results in interference.



(b) Spatial technologies allow time and frequency resources to be shared without interference.

Figure 2.2: Adding a spatial dimension.

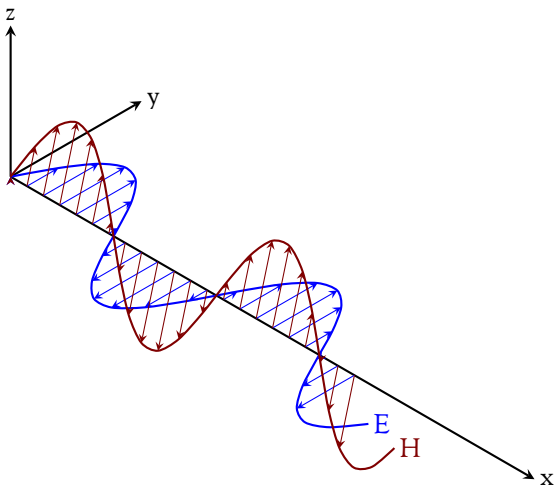


Figure 2.3: A linearly polarised electromagnetic wave propagating in the X direction. Derived form image by Cyril Langlois available at `http://www.texample.net/tikz/examples/polarizing-microscope/` and is available under a creative commons attribution licence `https://creativecommons.org/licenses/by/2.5/legalcode`.

2.2.2 Permeability, Permittivity and Conductivity

Permeability, μ, and permittivity, ϵ, are consonants in Maxwell's curl equations. They are properties of a medium in which an electromagnetic wave propagates. The unit of μ is Henrys per metre and ϵs units are Farads per metre. Equations 2.3 and 2.4 show how μ and ϵ can be expressed relative to their values in free space [7], where μ_0 and ϵ_0 are the values for free space with $\mu_0 = 4\pi \times 10^{-7} Hm^{-1}$ and $\epsilon_0 = 8.854 \times 10^{-12} Fm^{-1}$ and μ_r, ϵ_r are the relative permeability and permittivity respectively.

$$\mu = \mu_0 \mu_r. \tag{2.3}$$

$$\epsilon = \epsilon_0 \epsilon_r. \tag{2.4}$$

Conductivity, σ, is also a property of the particular medium. It is measured in Siemens per metre and is the reciprocal of resistivity. Permittivity can be expressed as a complex value with imaginary part equal to $\frac{\sigma}{\omega}$, where ω is the angular frequency of the wave.

These quantities help describe how common objects will interact with the propagating wave. Objects which have high conductivity, will attenuate penetrating electromagnetic waves significantly, whilst reflecting them efficiently. Knowledge of these properties, as well as the location and shape of various objects is essential for understanding the behaviour of a channel.

2.2.3 Electromagnetic Fields and Phasor Notation

It is mathematically helpful to express the phase and amplitude of an EM wave using complex values. As such, E and H can be expressed as in 2.5 and 2.6 [9].

$$\mathrm{E} = E_0 e^{[j(\omega t - kz) - az]} \hat{\mathrm{n}}. \tag{2.5}$$

$$\mathrm{H} = H_0 e^{[j(\omega t - kz) - az]} \hat{\mathrm{m}}. \tag{2.6}$$

Where $\omega = 2\pi f$ = angular frequency in $rads^{-1}$, t is time in seconds, k is wavenumber, z is distance propagated in metres and a is attenuation constant. Unit vectors $\hat{\mathrm{n}}$ and $\hat{\mathrm{m}}$ are mutually orthogonal with each other and the direction of propagation. The equations correspond to a wave with sinusoidal variation in amplitude based on both time and distance and exponential attenuation in amplitude dependent on distance. Where a medium is a good dielectric, i.e. a poor conductor, it is usually acceptable to remove the attenuation term [9]. It is also common to suppress the term $e^{j\omega t}$, therefore ignoring the time dependence.

2.2.4 Wave Impedance

Wave impedance, Z, is equal to the ratio of electric and magnetic field magnitudes, and can be expressed in Ohms. In a good dielectric, it is a real quantity. In a conductor, it becomes complex valued [9].

2.2.5 Phase Velocity

Phase velocity is the velocity of a point of constant phase on a wave. It is given in (2.7).

$$v = \frac{\omega}{k}, \tag{2.7}$$

where v is phase velocity, ω is angular frequency and k is wavenumber.

2.2.6 General Expressions for Wave Parameters

Materials can be described electromagnetically by μ, ϵ and σ. Using these constitutive parameters, the various parameters associated with a wave propagating in that material can be derived. Often simplifications can be made to the general expression if the material is a good conductor or dielectric. Table 2.1 shows the expressions for various quantities based on a materials constitutive parameters [2].

2.2.7 The Poynting Vector

The Poynting vector of an electromagnetic wave, S, describes both the magnitude and direction of power flow carried by the wave. Its magnitude is measured in Watts per square metre. The instantaneous value S is given by (2.8) [9].

$$\mathrm{S} = \mathrm{E} \times \mathrm{H}^*, \tag{2.8}$$

Parameter	Exact Expression
Attenuation constant $a\,[m^{-1}]$	$\omega\sqrt{\frac{\mu\epsilon}{2}[\sqrt{1+(\frac{\sigma}{\omega\epsilon})^2}-1]}$
Wavenumber $k\,[m^{-1}]$	$\omega\sqrt{\frac{\mu\epsilon}{2}[\sqrt{1+(\frac{\sigma}{\omega\epsilon})^2}+1]}$
Wave impedance $Z\,[\Omega]$	$\omega\sqrt{\frac{j\omega\mu}{\sigma+j\omega\epsilon}}$
Wavelength $\lambda\,[m]$	$\frac{2\pi}{k}$
Phase Velocity $v\,[ms^{-1}]$	$\frac{\omega}{k}$

Table 2.1: Electromagnetic Quantities

where E is the electric field vector and H is the magnetic field vector. S,E and H are mutually orthogonal and form a right hand set. Sometimes time average of power flow over one period of the wave is required in which case an expression is shown in 2.9.

$$\mathrm{S}_{av} = E_0 H_0 \hat{z}, \tag{2.9}$$

where E_0 is electric field amplitude and H_0 is magnetic field amplitude and $\hat{z}$ is a unit vector in the direction of propagation.

2.3 Antennas

Both antennas and propagation environment must be considered to successfully model a wireless communication channel [3]. For a system to radiate a propagating electromagnetic wave, a changing electric or magnetic field must be generated within a structure [2]. An antenna converts guided waves in a transmission line into waves that are radiated into 3D space where they continue to propagate or [9].

2.3.1 Input Response

Input response is a measure of how well matched an antenna is to its feed line. The reflection coefficient, which is denoted as S_{11} or Γ, describes the portion of an incumbent wave that is not radiated or absorbed into the antenna but is instead reflected back along the feed-line forming a standing wave [3].

The reflection coefficient has a real and an imaginary part. The magnitude usually quoted in dBs is referred to as return loss. A plot of return loss against frequency characterises an antenna's

input response, which can be used to evaluate bandwidth of resonances of the antenna. When calculating bandwidth a specific level should be quoted e.g. 10 dB bandwidth [3].

Voltage Standing Wave Ratio (VSWR), the ratio of maximum to minimum voltage on the feed-line, can also be used as a measure of antenna feed-line mismatch [3].

2.3.2 Near and Far Field

As a wave propagates away from the radiating structure it does so on a spherical wavefront. Close to an antenna the field patterns change rapidly with respect to distance and have both radiating and reactive energy [9]. The further a receiver is from the origin of the wavefront, the less reactive energy is present and the less spherical the incumbent wave appears to be. The distance at which reactive energy is negligible and at which the wavefront appears as planar rather than spherical by a receiving antenna is called the far field [2]. The distance d from the transmitting antenna at which the far field begins can be approximated by the rule of thumb: [5][9]

$$d > \frac{2L^2}{\lambda}, \tag{2.10}$$

where d is the distance from the antenna, L is the largest dimension of the antenna and λ is the wavelength of the radiation. This thesis will consider only radiation in the far field unless otherwise stated.

2.3.3 Polarisation

Where the propagating wavefront is planar (i.e. in the far field) the electric field E can be resolved into two components E^θ and E^ϕ which correspond to the θ and ϕ angles in a spherical coordinate system [9]. E^θ and E^ϕ are complex numbers with magnitude and phase. The magnetic field can be expressed in a similar way and is proportional to the electric field [3].

Polarisation describes the direction in which the electric field varies. For linearly polarised signals, the electric field varies in only a single direction perpendicular to the direction of propagation. Linearly polarised signals are often described as vertically or horizontally polarised with a single varying E^θ or E^ϕ component. It should be noted that horizontal/vertical alignment depends on the orientation of the coordinate system, and not all linearly polarised signals are vertically or horizontally aligned [3]. Figure 2.4 shows an instantaneous polarisation vector p. If p follows the path of the dashed red line with respect to time, then it is the polarisation vector of a linearly polarised wave that is not aligned with the vertical or horizontal directions.

Circular polarisations are also possible, described as left hand circular and right hand circular. Left hand circular means the field rotates anticlockwise, and right hand circular, clockwise, as the wave propagates away from the observation point. These can be related to E^θ and E^ϕ using Equations 2.11 and 2.12 [3].

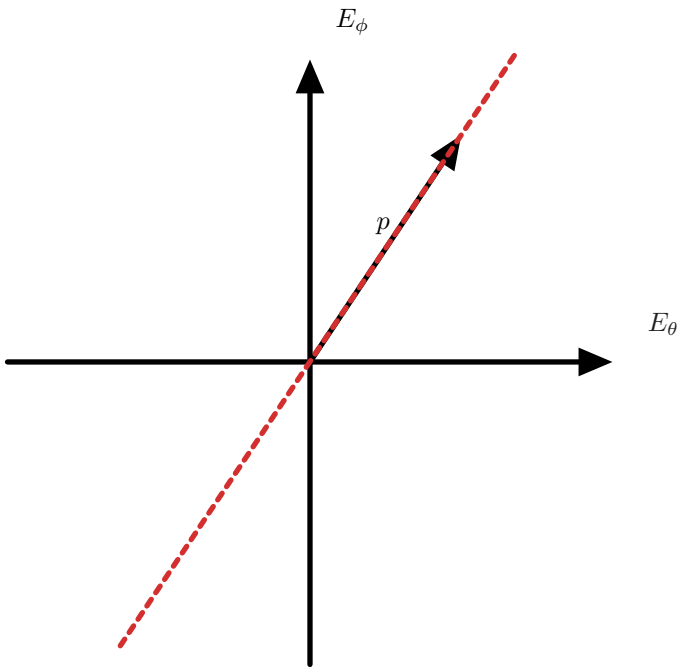


Figure 2.4: A polarisation vector.

$$E^{RHC} = \frac{1}{\sqrt{2}}|E^\theta + jE^\phi|. \tag{2.11}$$

$$E^{LHC} = \frac{1}{\sqrt{2}}|E^\theta - jE^\phi|. \tag{2.12}$$

More generally a signal will have elliptic polarisation which corresponds to variation in an elliptical pattern, with the major axis of the ellipse rotated at some angle to the vertical.

2.3.4 Directivity, Efficiency and Gain

Directivity is a measure of how much emitted radiation goes in one direction rather than another [3]. It is direction dependent. It can be expressed [9]

$$D(\theta,\phi,f) = \frac{\text{Radiation intensity in direction}(\theta,\phi,f)}{\text{Mean radiation intensity in all directions}} \tag{2.13}$$

$$= \frac{\text{Radiation intensity in direction}(\theta,\phi,f)}{\text{radiation intensity of isotropic source radiating same total power}}.$$

Directivity is usually quoted relative to an isotropic source in dBi, an isotropic source would have a directivity of 1 or 0dB. Directivity can be calculated with respect to either polarisation component,

E^θ or E^ϕ, or total power, including both polarisation components. In (2.14), it is shown how to calculate directivity for polarisation component E^θ [9]

$$D^\theta(\theta,\phi,f) = \frac{4\pi|E^\theta(\theta,\phi,f)|^2}{\oiint |E^\theta(\theta,\phi,f)|^2 + |E^\phi(\theta,\phi,f)|^2 \delta\Omega}. \tag{2.14}$$

Efficiency describes the amount of power that is radiated from the antenna relative to the total amount of power that the antenna sinks, including radiated power and any conductor and dielectric losses. Describing an antenna as having radiation resistance $R_R(f)$ and loss resistance $R_l(f)$ gives efficiency [9]

$$\eta(f) = \frac{R_R(f)}{R_R(f) + R_l(f)}. \tag{2.15}$$

Gain is the product of efficiency and directivity. In the same way as directivity, it is expressed for individual polarisations or total emitted power and is frequency dependent. An expression for θ polarised gain is shown below [9].

$$G^\theta(\theta,\phi,f) = \eta(f) D^\theta(\theta,\phi,f). \tag{2.16}$$

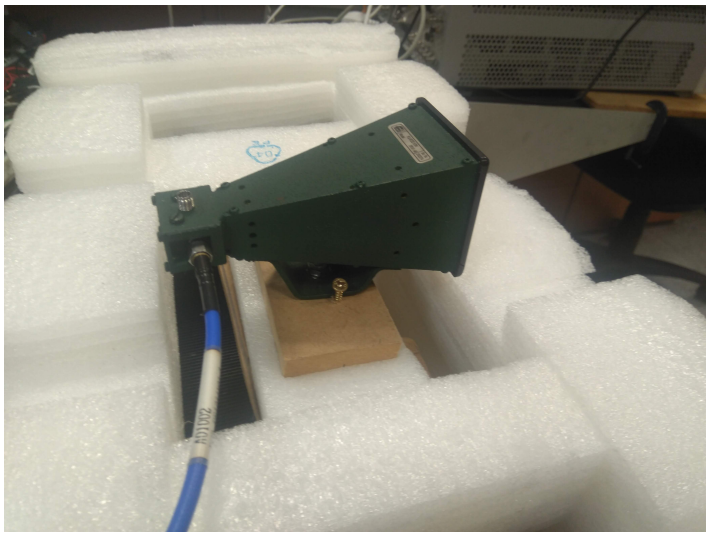
2.3.5 Antenna Effective Aperture

Antenna effective aperture is an area, oriented perpendicular to the direction of propagation, that would intercept the same amount of power from a radio wave as is produced at the terminals of the antenna in question. It is essentially a measure of how well an antenna is able to extract power from a wave with a given power density [9]. The effective aperture can have a clear physical interpretation in some cases; for example, the effective aperture of a dish antenna is closely related to the size of its dish. A less efficient antenna will have a smaller effective aperture. For an antenna with effective aperture $A_{eff}(m^2)$ receiving a wave of power density |S|(W/m^2). In (2.17), it is shown how power delivered to a load resistance, $P_r(W)$, is calculated [9].

$$P_r = A_{eff}|\mathrm{S}|. \tag{2.17}$$

The antenna effective aperture can be used to calculate the maximum gain of the antenna. An expression for maximum antenna gain, G, is given in (2.18) [2]. Figure 2.5 shows two antennas. Antennas will have different typical values for gain depending on their type. Gain for an efficient half wave dipole antenna is about 2.15 dBi. Other antennas may have much higher gains, for example a gain for a horn antenna might be 25 dBi and for a parabolic dish antenna 30-40 dBi.

$$G = A_{eff}\frac{4\pi}{\lambda^2}. \tag{2.18}$$



(a) A horn antenna.

(b) A parabolic dish antenna.

Figure 2.5: Antennas may assume various forms.

2.3.6 Frii's Path Loss Formula for Antennas in Free Space

Assume two antennas U and V are placed at a distance r away from one another such that both antennas are in each others' far field zones, and their directions of maximum gain and polarisations are matched. The incident radiation intensity |S| on antenna V as a result of power P_t being supplied to antenna U is given in (2.19) [9].

$$|\mathrm{S}| = \frac{P_t G_U}{area} = \frac{P_t G_U}{4\pi r^2}. \tag{2.19}$$

Using (2.17) the power delivered to the terminating resistance in antenna V is

$$P_r = \frac{P_t G_U A_{effV}}{4\pi r^2}. \tag{2.20}$$

And using the result from (2.18),

$$\frac{P_r}{P_t} = G_U G_V \left(\frac{\lambda}{4\pi r}\right)^2. \tag{2.21}$$

The formula (2.21) is Frii's transmission formula. It should be noted that received power is inversely proportional to r^2 in free space. Also note that (2.21) is only valid where polarisations are matched. Where there is a mismatch Frii's formula must be multiplied by the expression in (2.22) to account for this [10]. In (2.22), L is mismatch loss, p_w is the incident wave polarisation vector and p_a is the receiving antenna far field polarisation vector. $\cdot$ represents the vector dot product [9].

$$L = \frac{P_r}{P_m} = \left|\frac{p_w \cdot p_a^*}{|p_w| \times |p_a|}\right|^2. \tag{2.22}$$

2.4 Multipath

Multipath refers to the fact that in most real radio channels there are multiple paths between transmitter and receiver. These may be multiply reflected, scatted, diffracted or LoS. In general each multipath component can be described as a single phasor. These phasors are then combined at the receiving antenna. Assuming an isotropic point receiver, 2.23 gives the combined field E_R. Figure 2.6 shows how interaction with the environment leads to a multipath structure to the channel.

$$E_R = \sum_i A_i e^{j\theta}. \tag{2.23}$$

2.4.1 Fading

Fast Fading

Fast fading is caused by movement in the channel. The instantaneous field level may be determined by a coherent summation of multipath components at the receiving antenna. These components

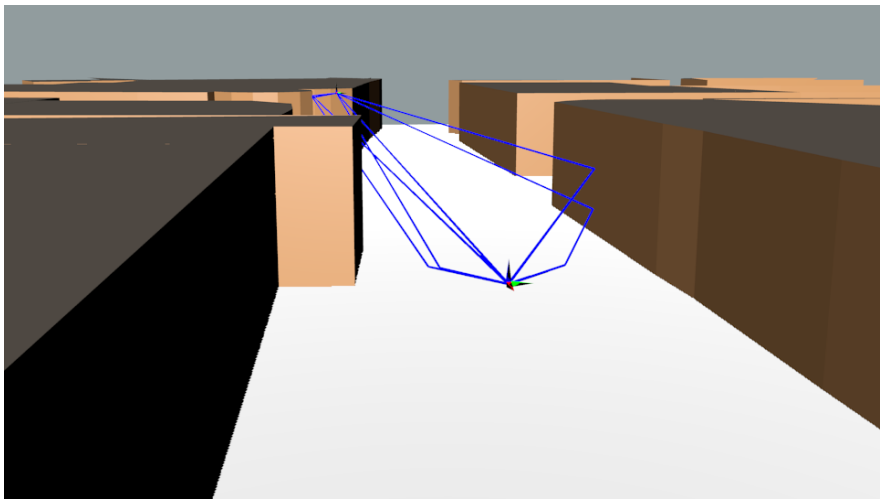


Figure 2.6: Multipath components in a multipath scenario. Contains information from open street maps, which is made available at `https://www.openstreetmap.org` under the Open Database License (ODbL) `https://opendatacommons.org/licenses/odbl/`.

may, at times, interfere constructively, and at others destructively. Because the phase relationships between multipath components will change completely due to motion in the order of a wavelength, this causes peaks and troughs in signal strength with a high spatial frequency. This is termed fast fading. Figure 2.7 shows how having multiple paths between transmitter and receiver can cause fades.

Doppler Shifts

Doppler shifts are generated due to the relative motion of radio terminals. For a path where the relative velocity of two communication terminals is Δv (ms^{-1}), the speed of the radio wave in the medium is c (ms^{-1}) and the transmission frequency is f_0 (Hz), the observed frequency f (Hz) is given by equation 2.24.

$$f = \left(1 + \frac{\Delta v}{c}\right) f_0. \tag{2.24}$$

In Fig. 2.10, a transmitter R is moving at speed u in the direction shown. For the stationary point N, the observed frequency will be higher than the emitted frequency. At the stationary point M the reverse is true.

Fig ?? shows the effect of varying the relative velocity between three different extremes at two different frequencies. Notice the change in scale of between the two subplots. At 3.5 GHz a relative velocity of 1 m/s causes a Doppler shift of approximately 10 Hz. At 60 GHz the Doppler shift for the same velocity is much greater, around 200 Hz. It can also be observed that Doppler shift increases linearly with relative velocity.

If two terminals are in a complex environment, there will be many propagation paths. When

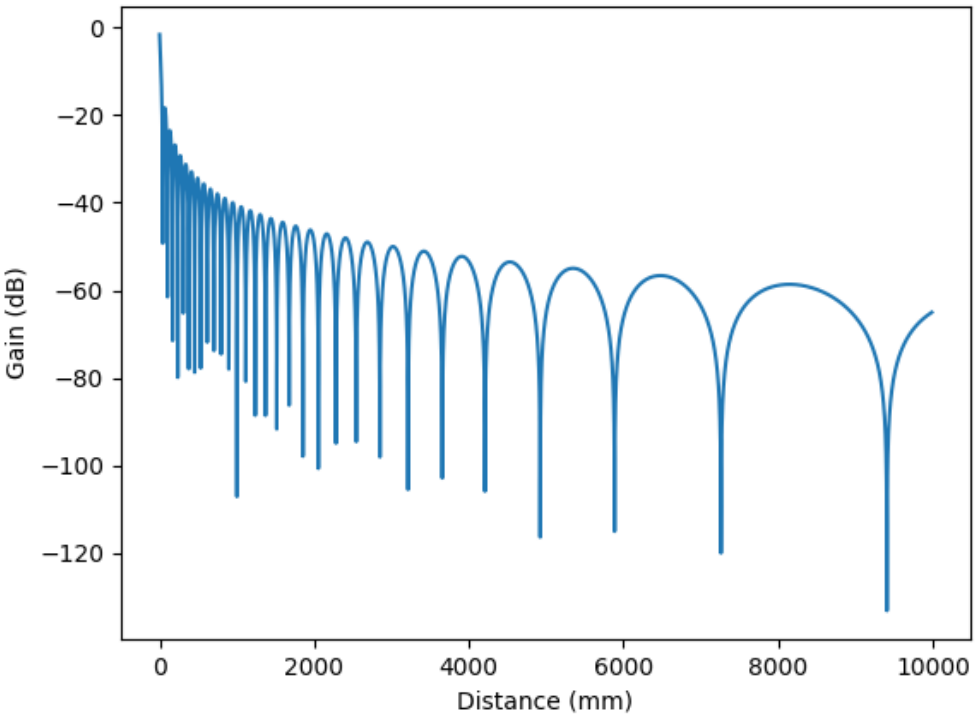


Figure 2.7: Fading caused by the interaction of a LoS path and ground bounce from 1 metre above the ground between two separated terminals.

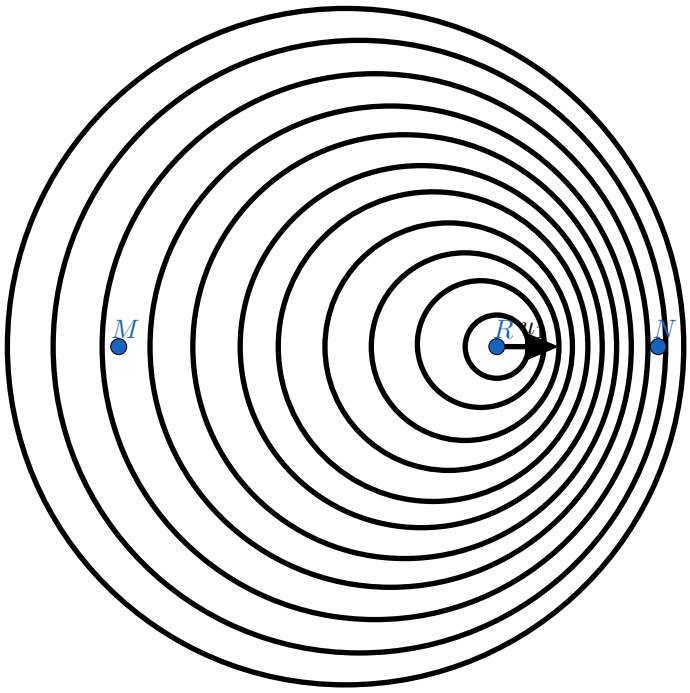
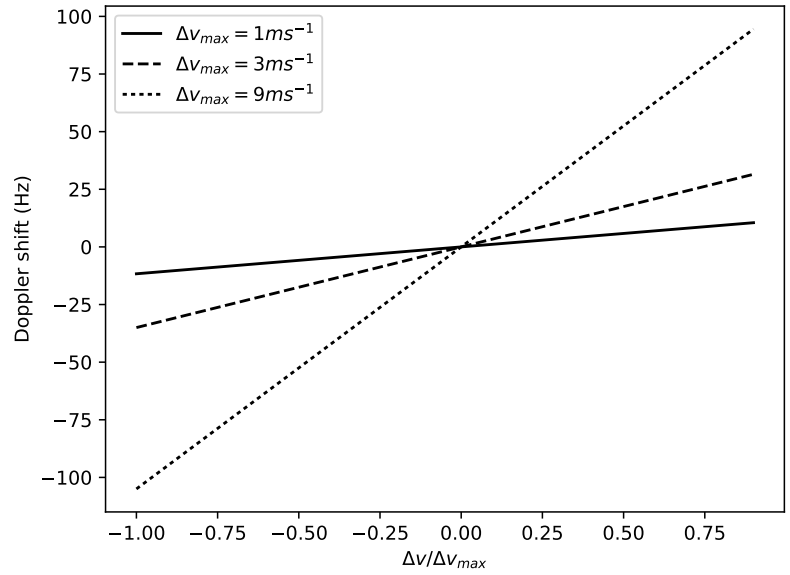


Figure 2.8: The Doppler effect

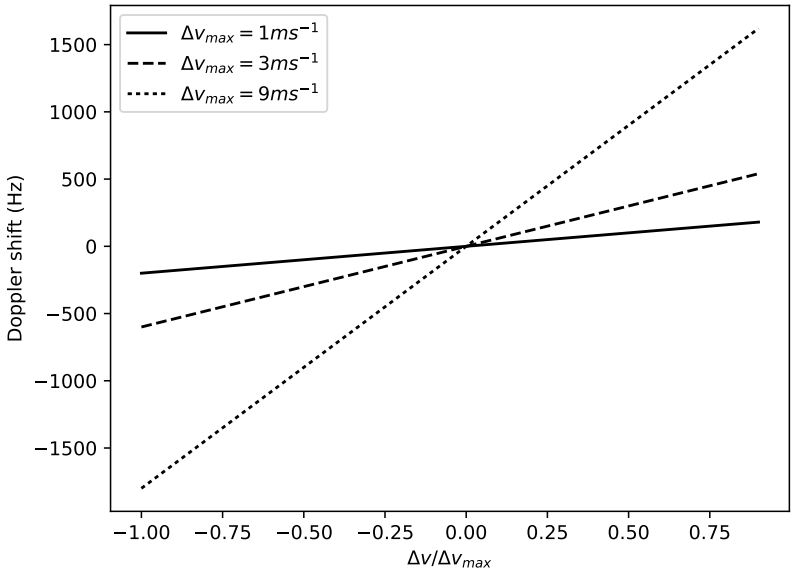
resolving their velocities along each path, the Doppler shift calculated for each path will, in general, be different. This results in a range of Doppler shifts or a Doppler spread. A Doppler power spectrum may be extracted. Fig 2.10 shows the Doppler power spectrum for a Rayleigh fading channel in which multipath components arrive from every azimuthal direction with equal probability. Knowledge of the Doppler shifts on multipath components allows the fast fading properties of the channel to be extrapolated.

Slow Fading

Slow fading refers to shadowing by large structures or terrain. This causes signal levels to drop as the radio waves do not penetrate such objects efficiently. By the time they reach the receiver they are heavily attenuated. This type of fading is called slow because it is consistent over an area several wavelengths in size. One may move an observation point several metres in the area behind a building and the fade experienced will be more or less consistent. This is shown in Fig. 2.11, where terminal S may move around the light grey shadowed zone and experience a relatively consistent fade. In the presence of only slow fading, the path loss may be approximated using a Gaussian random variable with mean and standard deviation in dB resulting in a log normal distribution of loss in Watts.



(a) Doppler shifts for various velocities at 3.5 GHz.



(b) Doppler shifts for various velocities at 60 GHz.

Figure 2.9: Doppler shift is proportional to velocity and carrier frequency.

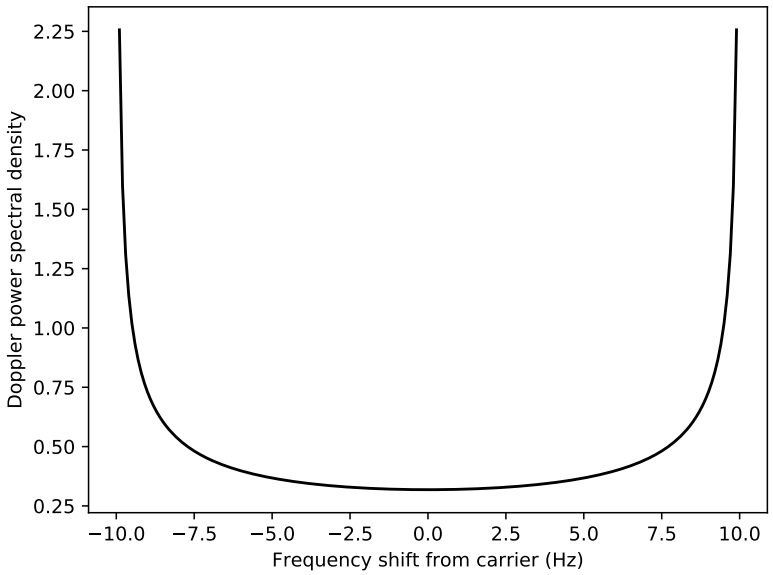


Figure 2.10: The Doppler power spectral density for a Rayleigh fading channel with equidistributed angles of arrival. The maximum Doppler shift is 10 Hz in this example.

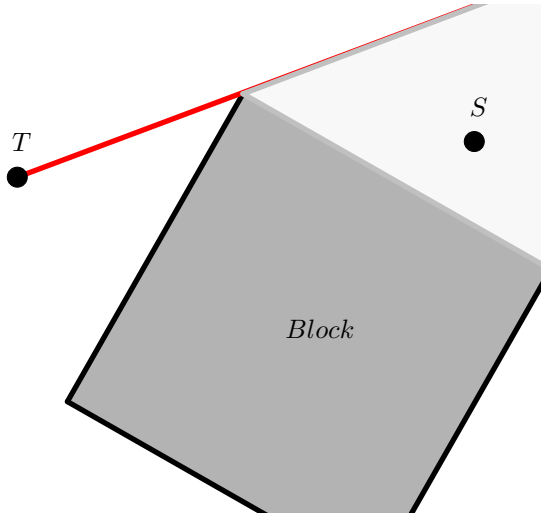


Figure 2.11: Slow fading

2.4.2 Delay

When an electromagnetic wave propagates, it does so at the speed of light. This means that there is a finite period of time between emission from the TX antenna and arriving at the RX antenna. Furthermore, usually there are many propagation paths between TX and RX. A LoS path may be present, as well as reflected diffracted and scattered paths. These paths have different lengths, and each introduce a different delay. Excess delay describes the delay between the first multipath component to arrive at the receiver and the n^{th} multipath component.

When ray paths are calculated in the ray launching tool, the lengths are stored. From this data, delay can be calculated for each multipath component. Mean excess delay, $\bar{\tau}$, is calculated as

$$\bar{\tau} = \frac{\sum^n (t_n - t_0)|V_n|^2}{\sum^n |V_n|^2}, \tag{2.25}$$

where t_0 is the time at which the first first multipath component arrives at the RX - and t_n is time taken for the n^{th} multipath component. V_n is the amplitude of the n^{th} ray. RMS delay spread is then given as

$$\tau_{rms} = \sqrt{\frac{\sum^n ((t_n - t_0) - \bar{\tau})^2 |V_n|^2}{\sum^n |V_n|^2}}. \tag{2.26}$$

Equations 2.25 and 2.26 are presented in discrete form because this is most natural when dealing with discrete ray paths. In practice, even measured data is often 'binned' into time slots, so the discrete from is again most appropriate. The coherence bandwidth, B (Hz), of the channel can be found from the τ_{rms} using

$$B = \frac{1}{\tau_{rms}}. \tag{2.27}$$

2.4.3 Angle of Arrival

One of the advantages of ray modelling is that spatial information is easily extracted. The angular spread of multipath components will have a significant effect on many aspects of the communications system, especially if beamforming is used.

Angles of arrival and departure can easily be extracted from ray tracing software. It is usually easiest to have some concept of a global spherical coordinate system and then, if necessary, resolve into a local spherical coordinate system for a rotated antenna using the global information. In this way, multiple antenna designs and rotations may be tested using a single shot of ray tracing.

2.4.4 Root Mean Square (RMS) Angular Spreads

RMS angular spreads, azimuth or elevation, are the power-weighted standard deviation of the azimuth or elevation direction of arrival.

Elevation

The mean elevation angle $\bar{\Phi}$ can be calculated from discete data using

$$\bar{\Phi} = \frac{\sum^n (\phi_n)|V_n|^2}{\sum^n |V_n|^2}. \tag{2.28}$$

From this r.m.s elevation spread is calculated using

$$\Phi_{rms} = \sqrt{\frac{\sum^n ((\phi_n) - \bar{\Phi})^2 |V_n|^2}{\sum^n |V_n|^2}}. \tag{2.29}$$

Azimuth

The mean azimuth angle $\bar{\Theta}$ can be calculated from discrete data using

$$\bar{\Theta} = \frac{\sum^n (\theta_n)|V_n|^2}{\sum^n |V_n|^2}. \tag{2.30}$$

From this r.m.s elevation spread is calculated using

$$\Theta_{rms} = \sqrt{\frac{\sum^n ((\phi_n) - \bar{\Theta})^2 |V_n|^2}{\sum^n |V_n|^2}}. \tag{2.31}$$

2.5 Modelling Antennas

Introduction

One of the challenges associated with ray tracing is how to include antenna radiation patterns in simulations. Since antennas are an important feature of the channel, it is a requirement of a good ray tracer to be able to import an antenna pattern and use it in signal strength calculations. The form of an antenna governs its electromagnetic behaviour. The most interesting property of the antenna, for propagation modelling purposes, is the far field radiation pattern. For any given direction it must be possible to return four complex antenna gain terms, corresponding to vertical to vertical, vertical to horizontal, horizontal to vertical and horizontal to horizontal gain for a propagating wave resolved into vertical and horizontal components. Due to limitations on memory in any real world machine, a continuous pattern cannot be stored. Instead a measured or simulated radiation pattern uses a certain angular resolution in the azimuth and elevation domains to sample a real world continuous radiation pattern.

2.5.1 Coordinate System and Polarisation Basis

It is important to discuss polarisation and coordinate systems. Several competing systems exist in the literature. Often these use similar or conflicting terms, meaning there is potentially unlimited opportunity for confusion. Here the meanings of the terms as used in this thesis are specified.

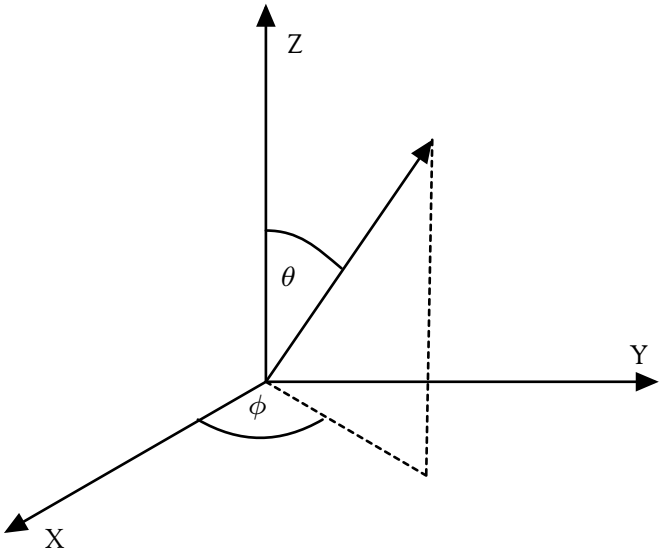


Figure 2.12: Coordinate system.

First a standard Cartesian coordinate system is used. A polar coordinate system is defined as shown in Fig 2.12. The polarisation basis is such that E_θ directions move through the pole at the Z axis, whereas E_ϕ move around it. For example, for a ray travelling along the X axis, E_θ would be in the Z direction and E_ϕ in the Y direction.

When the term 'co- polar' is used, it will always be in the context of an expected polarisation. For example, when feeding an antenna through a vertically polarised waveguide, the co-polar radiated field would be E_θ and the 'cross polar' E_ϕ. If the feeding waveguide were horizontal, them the reverse would be true.

2.5.2 Example Radiation Pattern File Format

The software developed in this thesis uses a '.cpat' format which is defined here for completeness. The antenna's radiation pattern is stored in two .cpat files. Each .cpat file is a series of lines with an arbitrary number of tab separated 'entries'. There are the same number of entries on each line. Each entry is four space delimited numbers. These are gain (linear) for the co-polar component, phase in radians for the co-polar component, gain (linear) for the cross polar component, and phase (radians) for the cross polar component. It is necessary to have two files, one for vertical illumination and one for horizontal illumination, giving all four complex components of gain.

```
elevation/altitude ("vertical")
angle [degrees from equator]
−90 −85 −75 ...          ... +75 +80 +85 +90
```

```
                −180 *   *   *                 *   *   *   * \n
azimuthal       −175 *   *   *                 *   *   *   * \n
  angle         ...
[degrees]       +180 *   *   *                 *   *   *   *,
```

Moving across a line from left to right, entries are equally spaced in the elevation domain from -90 degrees to 90 degrees. The lines in the file correspond to equal spacing in the azimuth domain from -180 degrees to 180 degrees. There is duplication of entries at the extremes of the file format. Note that the azimuth angle is identical to ϕ whereas the elevation angle is $90 - \theta$.

2.5.3 Rotating Antennas in Propagation Models

It is an attractive feature of a ray tool to be able to orient a particular antenna pattern in any direction within the environment. This allows realistic scenarios to be modelled, in which the coordinate system used to measure the antenna pattern is not necessarily exactly aligned with the coordinate system used in an environmental model.

A method for rotation of a point in a Cartesian coordinate system is multiplication by a 3x3 rotation matrix R. An expression for R is given in (2.32) [3].

$$\mathrm{R} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\omega_x & \sin\omega_x \\ 0 & -\sin\omega_x & \cos\omega_x \end{bmatrix} \begin{bmatrix} \cos\omega_y & 0 & -\sin\omega_y \\ 0 & 1 & 0 \\ \sin\omega_y & 0 & \cos\omega_y \end{bmatrix} \begin{bmatrix} \cos\omega_z & \sin\omega_z & 0 \\ -\sin\omega_z & \cos\omega_z & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{2.32}$$

In (2.32), ω_x, ω_y and ω_z correspond to rotations around x, y and z axis respectively, following a right hand rotation rule.

For a ray arriving at a receiver at angle of arrival specified as $(\theta^{AoA}, \phi^{AoA})$ relative to coordinate system originating at the receiver and having the same orientation as the global (environmental) coordinate system the direction vector of the arriving ray is given in Cartesian form as shown in (2.33) [3].

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \cos\phi^{AoA}\sin\theta^{AoA} \\ \sin\phi^{AoA}\sin\theta^{AoA} \\ \cos\theta^{AoA} \end{bmatrix}. \tag{2.33}$$

To calculate the direction vector of arrival in the coordinate system used to specify the antenna pattern (which will be rotated with respect to the global coordinate system) the rotation matrix R as shown in (2.34) [3] is used.

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \mathrm{R} \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \tag{2.34}$$

and converting into spherical coordinates with respect to the local coordinate system gives Equations 2.35 and 2.36.

$$\theta^{AoA'} = \tan^{-1}\left(\frac{\sqrt{x'^2 + y'^2}}{z'^2}\right). \tag{2.35}$$

$$\phi^{AoA'} = \tan^{-1}\left(\frac{y'}{x'}\right). \tag{2.36}$$

Next polarisation vectors must be considered. If the polarisation vectors relative to the global angle of arrival are given by $\mathfrak{p}^{\theta}$ and $\mathfrak{p}^{\phi}$ then the polarisation vectors relative to the local system will be in the same plane but out of alignment by an angle α with respect to $\mathrm{R}\mathfrak{p}^{\theta}$ and $\mathrm{R}\mathfrak{p}^{\phi}$ [3].

$$\mathrm{R}\mathfrak{p}^{\theta} = \mathrm{R}\begin{bmatrix}\cos\theta^{AoA}\cos\phi^{AoA}\\ \cos\theta^{AoA}\sin\phi^{AoA}\\ -sin\theta^{AoA}\end{bmatrix}. \tag{2.37}$$

$$\mathrm{R}\mathfrak{p}^{\phi} = \mathrm{R}\begin{bmatrix}-\sin\phi^{AoA}\\ \cos\phi^{AoA}\\ 0\end{bmatrix}. \tag{2.38}$$

$$\mathfrak{p}'^{\theta} = \begin{bmatrix}\cos\theta^{AoA'}\cos\phi^{AoA'}\\ \cos\theta^{AoA'}\sin\phi^{AoA'}\\ -sin\theta^{AoA'}\end{bmatrix}. \tag{2.39}$$

$$\mathfrak{p}'^{\phi} = \begin{bmatrix}-\sin\phi^{AoA'}\\ \cos\phi^{AoA'}\\ 0\end{bmatrix}. \tag{2.40}$$

$$\alpha = \cos^{-1}(\mathfrak{p}'^{\theta}\cdot\mathrm{R}\mathfrak{p}^{\theta}) = \cos^{-1}(\mathfrak{p}'^{\phi}\cdot\mathrm{R}\mathfrak{p}^{\phi}). \tag{2.41}$$

Using the results from Equations 2.35, 2.36 and 2.41 the far field radiation patterns (E^{θ}_{Rx} and E^{ϕ}_{Rx}) that correspond to a ray arriving at the receive antenna at angle of arrival (θ^{AoA}, ϕ^{AoA}) in the global coordinate system, can be expressed, as shown in (2.42) [3].

$$\begin{bmatrix}E^{\theta}_{Rx}\\ E^{\phi}_{Rx}\end{bmatrix} = \begin{bmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{bmatrix}\begin{bmatrix}E^{\theta}_{Rx}(\theta^{AoA'},\phi^{AoA'})\\ E^{\phi}_{Rx}(\theta^{AoA'},\phi^{AoA'})\end{bmatrix}. \tag{2.42}$$

As noted, antenna patterns are sampled at discrete angular intervals when measuring radiation patterns. This means a radiation pattern must be interpolated at angles that don't exactly align with one of the discrete entries. In the developed software, a simple bilinear interpolation is performed within the developed software to resolve this problem. This interpolation is performed for magnitude and phase.

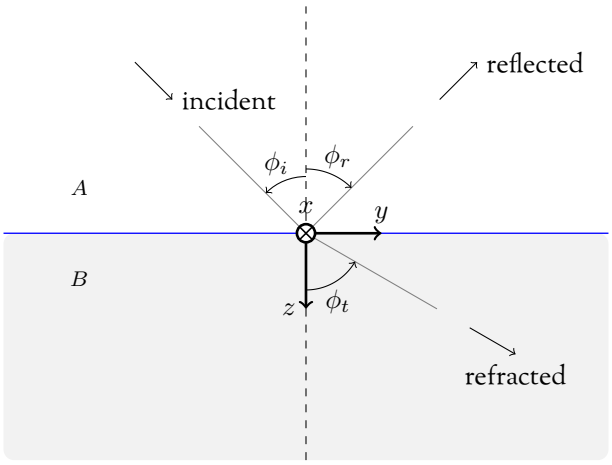


Figure 2.13: Incident, refracted and reflected waves at interface AB. Image is derivative work of original produced by Jimi Oke at `http://www.texample.net/tikz/examples/refraction/` and is available under a creative commons attribution licence `https://creativecommons.org/licenses/by/2.5/legalcode`.

2.6 Propagation by Reflection and Transmission

When a propagating wave, modelled as a ray, moves between two media (e.g. air and water), the ray interacts with the interface.

2.6.1 In Lossless Media

Consider Fig 2.13, it shows an incident ray at the interface between two media with different properties. The media are labelled A and B. The incident ray arrives at interface AB at the angle of incidence ϕ_i. When this ray interacts with interface AB, it is split into two further rays. The first is shown propagating away from the interface, back into medium A at angle ϕ_r, the angle of reflection. This is the reflected ray, and the reflection is described as specular when the interface is smooth. The other generated ray propagates into medium B at angle ϕ_t, the angle of transmission. This is the refracted ray. The energy contained in the incident ray is split between the reflected and the refracted rays. All three rays have their Poynting vectors in the scattering plane defined as the plane containing the incident Poynting vector and the normal to the interface.

Snell's law of reflection is given in (2.43). This is a simple expression based on Fermat's principle that the route taken by a ray between source and destination is the route with minimum traversal time (this principle is technically incomplete, but a useful conceptual model [1]). Considering a single reflection off an interface between two points in the same medium, it can be observed that equal incidence and reflection angles give the shortest distance reflected path and hence lowest traversal time.

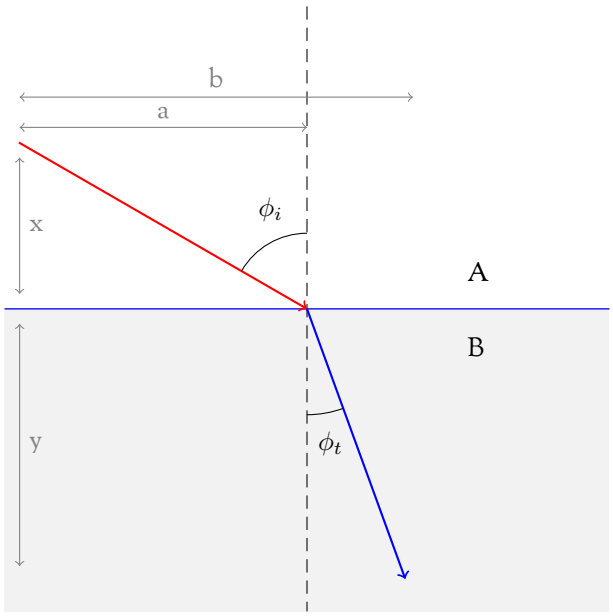


Figure 2.14: Refraction. Image is derivative work of original produced by Jimi Oke at `http://www.texample.net/tikz/examples/refraction/` and is available under a creative commons attribution licence `https://creativecommons.org/licenses/by/2.5/legalcode`.

$$\phi_i = \phi_r. \tag{2.43}$$

Fermat's principle is also useful when considering the ray that undergoes refraction, in which, the frequency of the wave is kept constant with respect to the incident wave, but the phase velocity and wavelength change. Fig 2.14 shows this.

Since the wavenumber of the two waves will be different, the traversal time t_{tot} to be minimised is [9]

$$t_{tot} = \frac{\sqrt{x^2 + a^2}}{v_A} + \frac{\sqrt{y^2 + (b-a)^2}}{v_B}, \tag{2.44}$$

where $v_{(A/B)}$ is propagation velocity through media (A/B) and all other symbols are as represented in Fig 2.14. Taking the derivative of (2.44) with respect to a gives (2.45).

$$\frac{\delta t}{\delta a} = \frac{a}{v_A\sqrt{x^2 + a^2}} + \frac{(b-a)}{v_B\sqrt{y^2 + (b-a)^2}}. \tag{2.45}$$

To find the minimum, (2.45) is equated to zero which gives [9]

$$0 = \frac{\sin(\phi_i)}{v_A} - \frac{\sin(\phi_t)}{v_B}. \tag{2.46}$$

$$\eta = \frac{c}{v} = \sqrt{\mu_r\epsilon_r} = \frac{\sqrt{\mu\epsilon}}{\sqrt{\mu_0\epsilon_0}}, \tag{2.47}$$

where η is refractive index, c is the speed of light in a vacuum and v is the velocity of the wave in the medium [9]. Using (2.47), this can be rearranged to produce (2.48), Snell's law of refraction.

$$\frac{\sin(\phi_i)}{\sin(\phi_t)} = \frac{\eta_t}{\eta_i}. \tag{2.48}$$

This is an expression for the angle at which the refracted ray propagates into an new material with a different refractive index. The field levels of the reflected and refracted waves are given by the Fresnel transmission and reflection coefficients. They will be different for the polarisations parallel and perpendicular to the scattering plane. Equations 2.49-2.52 show how reflection and transmission coefficients are calculated [9].

$$R_{||} = \frac{E_{r||}}{E_{i||}} = \frac{Z_B \cos\phi_t - Z_A \cos\phi_i}{Z_B \cos\phi_t + Z_A \cos\phi_i}. \tag{2.49}$$

$$R_{\perp} = \frac{E_{r\perp}}{E_{i\perp}} = \frac{Z_B \cos\phi_i - Z_A \cos\phi_t}{Z_B \cos\phi_i + Z_A \cos\phi_t}. \tag{2.50}$$

$$T_{||} = \frac{E_{r||}}{E_{i||}} = \frac{2Z_B \cos\phi_i}{Z_B \cos\phi_t + Z_A \cos\phi_i}. \tag{2.51}$$

$$T_{\perp} = \frac{E_{r\perp}}{E_{i\perp}} = \frac{2Z_B \cos\phi_i}{Z_B \cos\phi_i + Z_A \cos\phi_t}. \tag{2.52}$$

In Equations 2.49-2.52, $Z_{A/B}$ is wave impedance of media A/B. $\|$ denotes a parallel polarisation to, and $\perp$ denotes a perpendicular polarisation to, the scattering plane. Total reflected electric field can then be calculated as shown in (2.53).

$$\mathrm{E}_r = E_{r||}\hat{\mathrm{d}}_{||} + E_{r\perp}\hat{\mathrm{d}}_{\perp} = E_{i||}R_{||}\hat{\mathrm{d}}_{||} + E_{i\perp}R_{\perp}\hat{\mathrm{d}}_{\perp}. \tag{2.53}$$

In (2.53), $\hat{\mathrm{d}}_{\|}$ and $\hat{\mathrm{d}}_{\perp}$ are unit vectors parallel and normal the scattering plain respectively. The incident electric field may take any polarisation state and can be expressed as in (2.54).

$$\mathrm{E}_i = E_{i||}\hat{\mathrm{d}}_{||} + E_{i\perp}\hat{\mathrm{d}}_{\perp}. \tag{2.54}$$

It is useful to express these calculations in matrix form. Commonly calculation of the reflected field is conducted using (2.55) [9].

$$\mathrm{E}_r = R\mathrm{E}_i, \tag{2.55}$$

where,

$$\mathrm{E}_r = \begin{bmatrix} E_{r||} \\ E_{r\perp} \end{bmatrix}, \mathrm{E}_i = \begin{bmatrix} E_{i||} \\ E_{i\perp} \end{bmatrix}, R = \begin{bmatrix} R_{||} & 0 \\ 0 & R_{\perp} \end{bmatrix}. \tag{2.56}$$

Similarly the refracted field is calculated using

$$\mathrm{E}_t = T\mathrm{E}_i \tag{2.57}$$

where,

$$\mathrm{E}_t = \begin{bmatrix} E_{t||} \\ E_{t\perp} \end{bmatrix}, \mathrm{E}_i = \begin{bmatrix} E_{i||} \\ E_{i\perp} \end{bmatrix}, T = \begin{bmatrix} T_{||} & 0 \\ 0 & T_{\perp} \end{bmatrix}. \tag{2.58}$$

2.6.2 In Lossy Media

In lossy media, Snell's law of refraction, (2.48), does not hold. Snell's law of reflection still holds as do Equations 2.49-2.58, assuming wave impedance Z is understood to be complex [9]. An expression for Z is shown in (2.59) where σ is conductivity in Siemens per metre [9].

$$Z = \sqrt{\frac{j\omega\mu}{\sigma + j\omega\epsilon}}. \tag{2.59}$$

To address the problems associated with the apparent angle of transmission not being that described by Snell's law of refraction, [6] presents a series of equations for the apparent transmission angle based on the material properties of a non lossy to a lossy dialectic interface and [4] and [8] describe how to model uniform and non-uniform waves impinging on an interface between two different lossy materials respectively. However, the technically correct handling of lossy transmission has not been incorporated into software described later in the thesis. The justification for this is that if the material has a significant conductivity, i.e. is lossy, then the refracted ray is likely to attenuate quickly to an extremely low field level, and thus these channel components will generally lack significance.

2.7 Materials Characterisation

2.7.1 Introduction

For use in ray tracing and other electromagnetic modelling activities it is desirable to be able to determine the electromagnetic properties of materials. Permittivity and permeability are most commonly required.

2.7.2 Experiment

To measure the permittivity of various materials, a Keysight E4991B Impedance Analyzer was used. Small discs of each material were cut and, if necessary, manipulated to obtain a smooth finish (when possible, not for carpet). The thickness of each sample was measured. The samples were loaded into the device and data was recorded to memory for analysis. The E4991B Impedance Analyzer is shown in Fig 2.15. A piece of calibration material, having near constant permittivity over the measured frequency range, was loaded in to the device. The result of this calibration is shown in Fig 2.16.

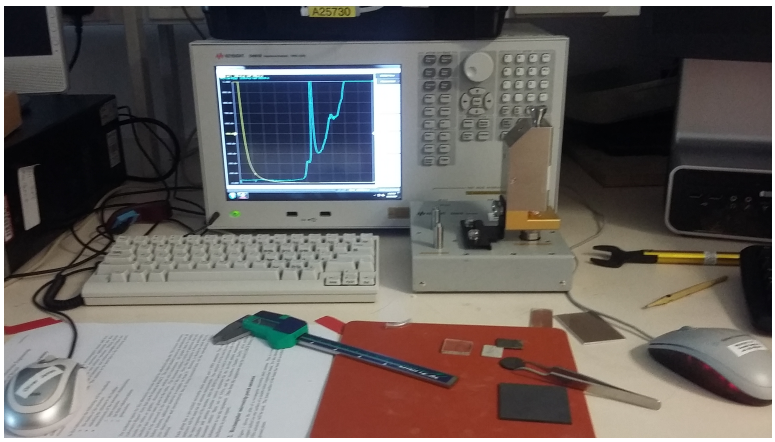


Figure 2.15: Experiment

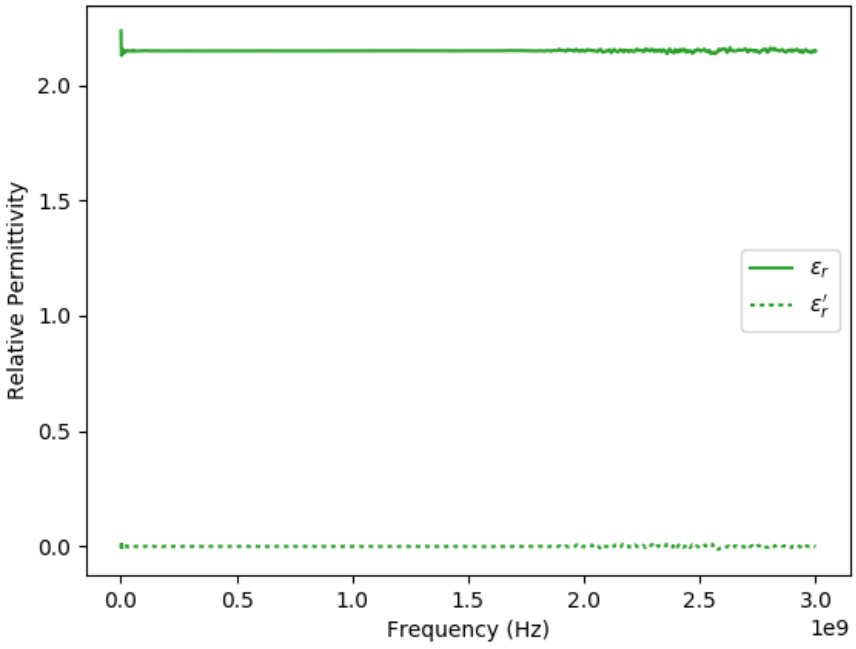


Figure 2.16: Calibration

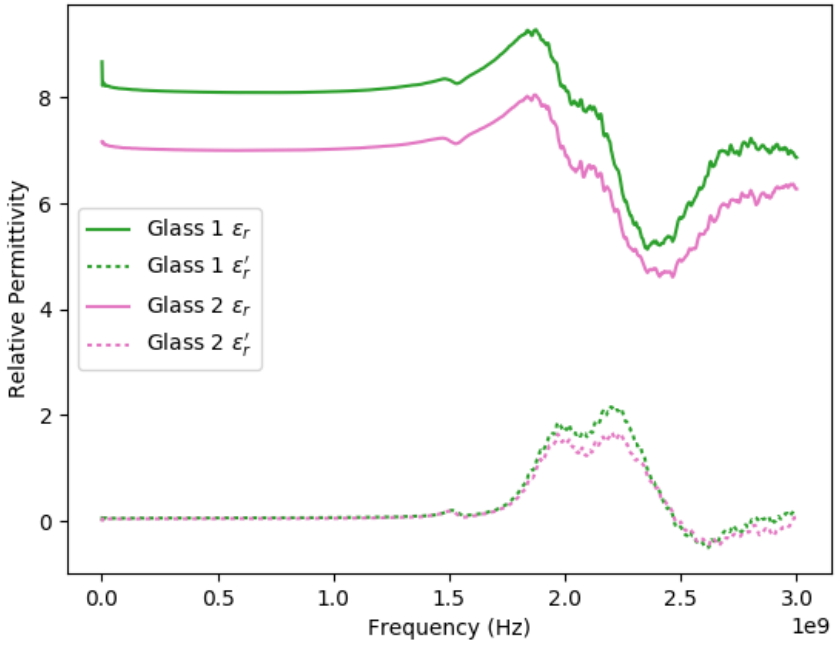


Figure 2.17: Glass

2.7.3 Materials Classified

Values of permittivity were obtained for: ceramic tile, softwood (unknown), softwood endgrain (pine), softwood crossgrain (pine), PVC, polypropylene, poly-carbonate, plywood, plasterboard, hardwood endgrain (oak), hardwood crossgrain (oak), stone, MDF, hardwood crossgrain (mahogany), hardwood endgrain (mahogany), linoleum, leather, insulation, green hardwood (apple), glass 1, glass 2, firebrick, concrete, carpet tile, brick and acrylic from 1 - 3000 MHz.

2.7.4 Results Discussion

A number of interesting trends were found. In Fig 2.17, it can be seen that two samples of glass, taken form different sources, have different permittivity values. Both samples had a minima in permittivity at about 2.4 GHz, but a difference in permittivity, regardless of frequency, shows that different manufacturing processes or chemical makeup can affect permittivity in apparently similar materials. In Fig 2.18 green wood, recently cut from an apple tree is compared with seasoned oak. It can be observed that the real part of permittivity is much higher for the green wood, presumably due to its higher water content. This trend is observable when comparing the green wood to any of the other samples of seasoned wood. This suggest if trees or plants are to be modelled explicitly this must be taken into account.

In Fig 2.19, it is shown that for materials that have a directional structure, e.g. wood, where the

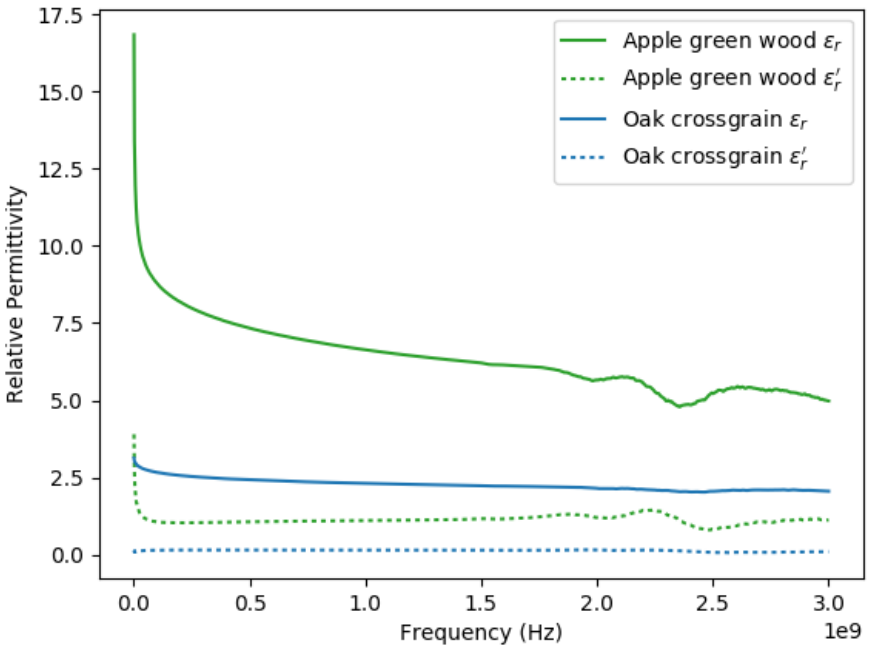


Figure 2.18: Green wood

grain has a clear direction, the orientation of the material may affect its electromagnetic properties. For all samples of wood where a crossgrain and endgrain sample could be compared, permittivity was greater for the endgrain sample. Although the sample size was clearly limited, it is interesting to note that this effect was least pronounced for the softwood sample where the grain structure was the least dense.

The information gathered in this chapter was used later in this thesis when two new ray tools were developed. The electromagnetic theory presented here was incorporated into the ray modelling tools presented and used in chapters 4, 5 and 6. Materials parameters were used to assign electromagnetic properties to virtual environments, making simulations more realistic. The ray tracing tools in subsequent chapters use the plane wave assumption and model polarisation using two principal components. The assumption was made that materials would retain approximately the same permittivity across the range of frequencies of interest.

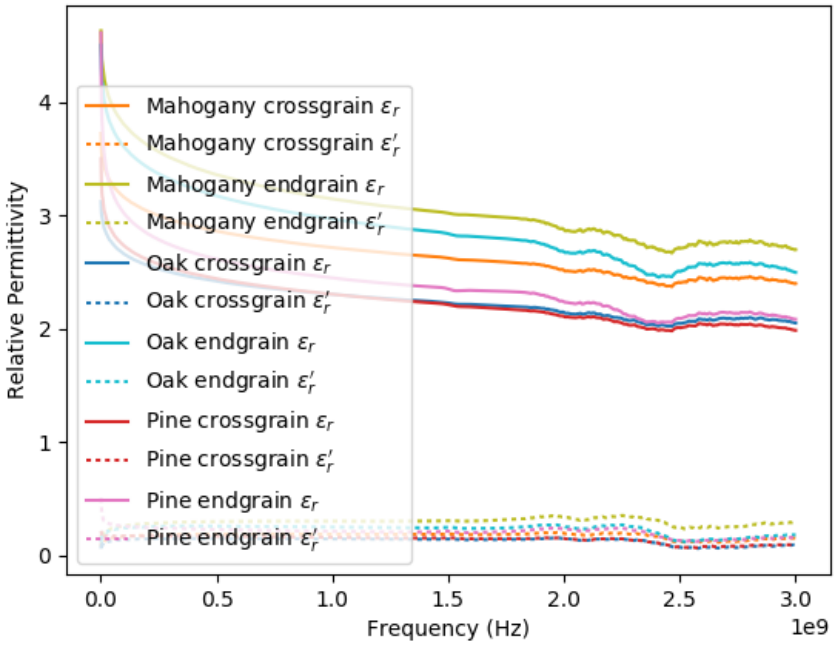


Figure 2.19: Grain direction

Bibliography

[1] The feynman lectures on physics vol. i ch. 26: Optics: The principle of least time. `http://www.feynmanlectures.caltech.edu/I_26.html`. (Accessed on 07/21/2016).

[2] Constantine A Balanis. *Antenna theory: analysis and design*. John Wiley & Sons, 2016.

[3] Mellios Evangelos. *Physical Layer Analysis of Communication Systems Integrating Antenna Characteristics and Propagation Modelling for Indoor and Outdoor Environments*. PhD thesis, University of Bristol, 2013.

[4] J. Holmes and C. Balanis. Refraction of a uniform plane wave incident on a plane boundary between two lossy media. *IEEE Transactions on Antennas and Propagation*, 26(5):738–741, Sep 1978.

[5] J.D. Kraus. *Antennas*. McGraw-Hill, 1988.

[6] M. C. Lawton and J. P. McGeehan. The application of a deterministic ray launching algorithm for the prediction of radio channel characteristics in small-cell environments. *IEEE Transactions on Vehicular Technology*, 43(4):955–969, Nov 1994.

[7] J Clerk Maxwell. A dynamical theory of the electromagnetic field. *Philosophical transactions of the Royal Society of London*, 155:459–512, 1865.

[8] R. D. Radcliff and C. A. Balanis. Modified propagation constants for nonuniform plane wave transmission through conducting media. *IEEE Transactions on Geoscience and Remote Sensing*, GE-20(3):408–411, July 1982.

[9] Simon R. Saunders. *Antennas and Propagation for Wireless Communication Systems*. Wiley, 1999.

[10] Warren L. Stutzman and Gary A. Thiele. *Antenna Theory and Design*. Wiley, 2012.

Chapter 3

Propagation Modelling

Some results in this chapter were obtained in collaboration with colleagues at the University of Bristol. Specifically, section 3.2 shows results from the joint work [14]. The antenna design under test, ray tracing results, Fig 2.12 and antenna radiation pattern measurements were performed by colleagues while the remainder of the analysis of the results and and the proposed antenna evaluation method were the author's own work.

3.1 Propagation Modelling Review

There are a variety of reasons why propagation modelling may be attempted and there are many different metrics that describe a channel that may be extracted from the modelling process. Depending upon the particular engineering problem, it is essential to pick a model with appropriate qualities. The ideal model is quick to run, accurate, precise and easy to initialise. In reality, it is extremely difficult to meet all these requirements at the same time. It is usually known in advance what information is of interest, whether that is field levels across an antenna array, delay spread in a city or coherence time for a mobile channel. This means that propagation models can be tailored to specific applications. Relaxing requirements in areas of no particular interest means models can be optimised to retrieve the desired information in the most efficient way for a stated problem.

3.1.1 Statistical Methods

Statistical models are based on statistical distributions derived from real world measurement campaigns. Statistical models may be classified as geometry based stochastic models or non-geometric stochastic models; this depending on whether some initial deterministic information about the

layout of the environment is provided. Various statistical methods were adopted for use in planning cellular and broadcast networks. Their simplicity and sufficiency for cellular planning are key strengths. The Lee model, COST 231 Walfisch-Ikegami model and Okumura-Hata model are examples in this vein [18][27][40]. Many others exist [15].

The main advantage of statistical models is that they are simple to implement and tractable, working on a few basic input parameters. However, these models do not take into account much of the site-specific information for any given environment, and are therefore merely representative of a set of similar environments, often tagged with descriptions such as 'rural' or 'urban'. Often analysis is restricted to the azimuth plane [15]. Important deterministic information can be left out such as material properties and in the case of non-geometric models layout information. Even if geometric models allow a basic layout to be entered, it is often simplified. Work has been undertaken in order to develop statistical methods that produce both timing and angle of arrival predictions, which are difficult to extract from a simple expression [47]. When using statistical methods, predictions can only be interpreted as broadly representative when the model is run for an environment significantly different to the one in which the supporting measurements were taken [31]. Statistical input data must be obtained for each frequency under test. This means for newer technologies further measurements would have to be obtained which is likely to beyond the means of many stakeholders [31].

3.1.2 Finite Difference Time Domain (FDTD) and Method of Moments (MoM)

FDTD models mimic electromagnetic interactions closely [55], however they are computationally unfeasible for room-sized three-dimensional environments [44]. The finite domains that an environment is divided into must be smaller than both the wavelength of interest and the smallest dimension any object in an environment to give meaningful results. Some calculation must be performed for each domain, which rapidly becomes intractable as room size is increased. Currently, only environments with a size of the order of ten wavelengths can be reasonably modelled using this method [57]. Work to decrease the computational requirements e.g. [1][9] has been performed, but has been insufficient to achieve timely modelling at large scales. The FDTD technique may be useful when augmenting a hybrid model for large scale simulations, indeed results from small-scale FDTD models can be incorporated into a more abstracted model as in [54].

Whilst FDTD is a differential equation solving method, the MoM solves integral equations. It is a boundary element method, meaning that again a discretisation is necessary, and suffers from the same tractability problems for large environments.

3.1.3 Ray Modelling

Ray modelling is a propagation modelling technique in which the travelling wave-front of a radio wave is represented as a series of rays that obey the laws of geometrical optics. Ray models take as an input a computer model of an environment. In this way, the interaction of radio wave and environment is captured, making results more accurate for a specific environment than statistical methods, whilst being very much less computationally intensive than the FDTD or MoM methods for a large environment.

Outdoor vs Indoor Ray Modelling

Ray models are often designed specifically for either indoor or outdoor environments. Those designed for outdoor use usually have lower granularity and are more scaleable, whereas those for indoor use model the environment more precisely for improved accuracy [25][26][42]. Outdoor ray models are often forced to make approximations to simplify modelling. For example, an outdoor ray model may perform path finding in a 2D plane, only using building height data in electromagnetic models [3], whereas modern indoor models might perform full ray tracing in three dimensions since this is much more tractable on the much smaller scales involved. It is also easier to create a full 3D model of an indoor space, again because of its smaller size and availability of architectural plans [5], although, increasingly, 3D data for whole cities and landscapes is available.

Ray techniques can also be used to investigate outdoor to indoor and indoor to outdoor propagation. One author presents some of the complexities of this, whilst also demonstrating that ray tracing can achieve comparable performance to the FDTD technique in outdoor to indoor modelling [57]. Another identifies one of the problems with combining indoor and outdoor models, which is that as well as a low fidelity model of a city one also needs an extremely high fidelity model of a building in order to accurately predict the indoor propagation [7].

Shooting or Tracing

Ray models may be classified as either a ray shooter or a ray tracer.

Ray Shooting Ray shooting either uses a pin cushion method, or environment driven method. When using the pin cushion method, rays are fired out as evenly as possible around a TX location [20][32][43]. After propagating through the environment, the launched rays will not exactly intersect the receiver point. Fig. 3.1 shows the most common method by which it is decided if a ray is received. The radius of the reception sphere is based on the angle between cast rays and path length to the receiver [21][45][3]. Alternatively, the arriving ray is checked to be within angular bounds created by a virtual transmitter [5], as shown in Fig 3.2. The problem of how to evenly distribute fired rays around the transmitter reduces to distributing points as evenly as possible on the surface of a

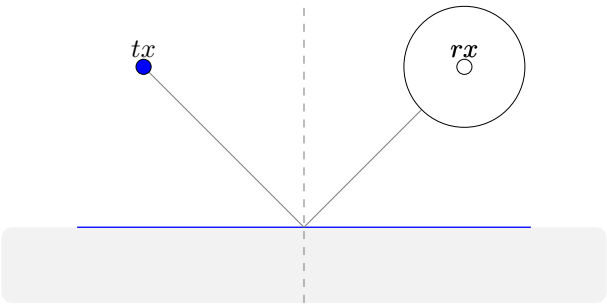


Figure 3.1: Radius. Image is derivative work of original produced by Jimi Oke at `http://www.texample.net/tikz/examples/refraction/` and is available under a creative commons attribution licence `https://creativecommons.org/licenses/by/2.5/legalcode`.

sphere. Having found a solution, ray directions can then be found by constructing a vector from the centre of the sphere to the points. No exact solutions exist for the number of points required to achieve satisfactory resolution in a ray tracer. In fact, twelve is the maximum number of rays that would allow an exact solution. The usual method used in pin cushion ray launching applications is to construct a regular icosahedron whose vertices lie on the unit sphere. Each face is subdivided into more triangles and then the vertices are projected onto the unit sphere [46][45][32]. Whilst this is not an exact solution, in practical terms, it is sufficient for accurate ray tracing.

Each ray covers a spherical cap on the spherical wavefront. To make sure that the whole wavefront is represented, these spherical sections must slightly overlap. This means no propagation paths are missed. However, because of this overlapping, there may be double counts of duplicate propagation paths. A post-processing step in which duplicate rays are filtered is required [5].

In an environment driven mode, rays are fired to the centre of tiles covering the faces in the environment. This has the advantage that no useless rays that hit nothing are fired. Also divergence is less severe as an appropriate spatial resolution can be defined by picking tile size appropriately. Rays from the tiles can be fired directly at a receiver so the method for finding if a ray has been received is easier. However, because these rays may not be specular, it is more difficult to model a pure specular reflection.

There are cases where ray launching is unnecessarily computationally expensive, with simple environments being the worst case [31][39]. It also has poorer accuracy than image based techniques, especially over long distances due to divergence [3][26][31]. For complex environments it has been successfully used in [32][33][34][5]. In general, the more complex the environment, the more attractive ray launching becomes as it scales more favourably with the maximum allowed number of reflections than the image method, which is introduced next.

The Image Method The image method works out direct paths between an infinitesimal TX and RX by creating virtual images of the TX through the reflecting facets of the room [53][49][30][13][52][29][19].

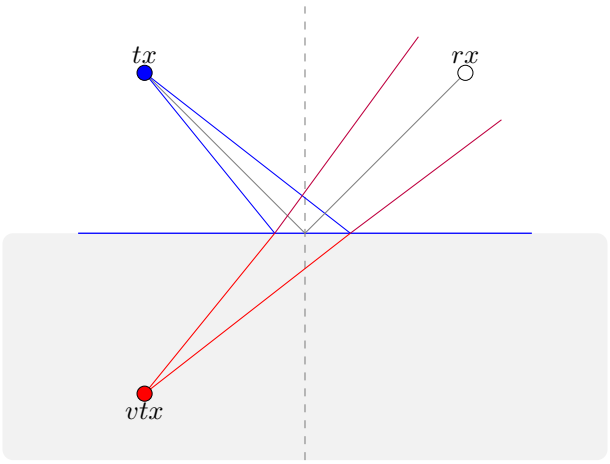


Figure 3.2: Angular Bounds. Image is derivative work of original produced by Jimi Oke at `http://www.texample.net/tikz/examples/refraction/` and is available under a creative commons attribution licence `https://creativecommons.org/licenses/by/2.5/legalcode`.

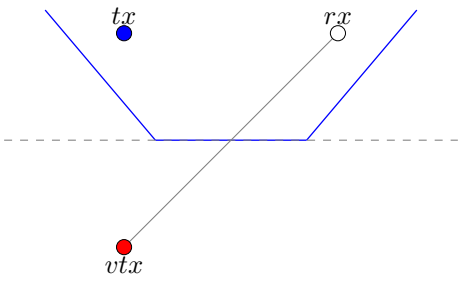


Figure 3.3: Image method

As there may be an infinite number of such paths the maximum number of allowed reflections is specified in order to truncate the model. A path is valid if there is a straight line between the source and an n^{th} order image of a receiver which intersects the n facets used to generate the image but no other facets. Figure 3.3 shows a first order reflection through a wall resulting in a ray path as identified by the image method.

Other phenomena, such as refraction, may also be accounted for using the image method by altering how images are generated. When modelling a large number of interactions the method becomes cumbersome due to the exponentially increasing number of images that must be considered [45][26]. In an environment with many facets, where at each facet refracted, reflected and scattered rays are generated, the method becomes unattractive unless suitable optimisation techniques are used. The advantages of imaging are that there are no resolution problems contributing twice counted or missing ray paths, and no divergence errors in electromagnetic calculations. In [52], it is also suggested that it is easier to include diffraction in an image based ray engine.

Hybrid Method Hybrid models use a combination of imaging and ray shooting [51]. A ray shooting method is conducted as normal and then, since all reflecting walls on a particular path are now known, the image method is used to remove inaccuracies generated by the ray shooting method at a fraction of the computational cost of a full image search. It is also possible to use the image method for some propagation phenomena, and the ray shooting method for others, as in [16].

3.1.4 Ray Modelling for 5G

Whilst ray modelling has been used for propagation modelling since the early nineties, simpler statistical methods have often been preferred for coverage predictions. However, for newly proposed mmWave links, there are number of propagation phenomena which make ray modelling more desirable. The first is the low penetration of mmWaves. This means poor data on the internal structure of walls, attributed to the difficulty on discovering this structure nondestructively, is less of a problem because refracted rays are less significant. Diffraction is also less significant, which, when modelled, contributes significantly to total simulation time because finding diffracted paths is computationally expensive. Recent developments in computing technology mean ray tracing is more practical. The data made available by ray modelling, specifically the 3D spatial and poliametric information, are more useful, because beamforming techniques are required to mitigate the relatively higher propagation losses experienced in mmWave systems, and these need directional data to be properly evaluated [10].

A requirement for millimetre wavelength ray tracing is to be able to include the effects of diffuse scattering. There has been some debate in the literature about the importance of diffuse scattering, it generally seems that for scenarios where a transceiver is close to a rough surface it is an important propagation mechanism [16]. Directional data is also essential for simulation of MIMO links. A short survey of commercial and academic tools is presented here.

University of Bologna Ray Tracers

In [12], the University of Bologna ray tracer is described. The ray tracer follows a combined imaging and ray launching technique, using a visibility matrix in order to reduce computation time. The ray tracer does include scattering using the effective roughness (ER) model [11]. The main disadvantage is the need to tune the ER model using the scattering parameter S, which is usually done from measurements. The model must be re-parameterised for new frequencies. The tool is extensively optimised, and has been used in several papers.

The Bologna team also have a specialist indoor tool. It appears that this tool is able to model environments in greater detail. It is used in [16].

University of Bristol Ray Tracers

ProPhecy is the mature outdoor ray tracing tool initially developed at the University of Bristol [41][50]. The tool was highly optimised and includes several propagation mechanisms. Also developed at the university was a tool specifically for indoor use, as used in [14]. The code base for the indoor version was unfortunately, lost. The outdoor version is in continued development, but currently has some limitations due to the various optimisation strategies used.

Remcom Wireless InSite

Wireless InSite is the name given to Remcom's ray tracer [2]. It includes Lambertian end ER scattering models. The tool has been tested against the mmWave measurement campaign in [38] and this gave good results. According to a webinar hosted by Remcom, the product is able to model scattering after any number of reflections, and allows the amount of reflection to be limited to reduce computation time.

Altair WinProp

In [22], an Altair propagation modelling tool is described. The model is fully 3D and benefits from techniques such as preprocessing the input database and finding only dominant paths. Diffuse scatter is not mentioned in documentation.

Siradel Volcano/S_5GChannel

Volcano/S_5GChannel are the commercial names of propagation modelling tools from Siradel [8]. The 5G specific aspects are some of the most advanced of any of the investigated tools, with LIDAR enhanced modelling for enhanced simulation of rough surfaces and vegetation [48]. MIMO studies have also been performed [4]. The tool is fully 3D and allows attractive plotting of results in CAD software.

Dublin City University Ray Tracer

A 2.5D ray tracer was developed at Dublin city university [23]. The tool boasts excellent performance when compared to the COST 231 model, and in [23][24] optimisation techniques are presented, which are grid based. This achieves fast execution speed for urban coverage prediction, which appears to be the tools intended use.

3.2 Case Study – An Antenna Evaluation Method Using Ray Tracing

In this section, it is demonstrated that ray modelling can be used to solve engineering problems in the real world. This is illustrated using a case study in which modelling is incorporated into an antenna design and test process. This method was the subject of [14].

3.2.1 Background

Antennas and radio wave propagation constitute the basic elements of the wireless channel. They, therefore, determine the quality and the reliability of the wireless link and hence have a great impact on the QoS offered by a whole system. Modelling of both these aspects of the channel is of interest to device designers. In a body-centric wireless network, the human body imposes several extra challenges for propagation modelling and antenna testing, such as electromagnetic absorption, propagation shadowing and polarisation misalignment [6][17][56][37][28].

This work contributes a method by which the performance of a wearable antenna in a BLE in-home network can be quantified - ensuring that high link quality is provided for body-centric applications. The method combines antenna characteristics with a physical model of the wireless channel. The communication system performance is evaluated for various antenna orientations and locations within a residential environment. In contrast, most methods in literature are constrained to either a particular room or specific link type. Physical modelling of the channel is performed using ray-tracing simulations. The ray-tracer provides the multipath propagation characteristics of the channel in a residential setting in which body-centric networks are likely to operate.

An on-body antenna design was proposed and used as a candidate for evaluation of the proposed method. The antenna was a patch antenna which provided radiation pattern switching across the 2.4 GHz operating band without an additional switching mechanism. This was achieved by merging an omnidirectional and a directional radiation mode of the antenna together within a single band, in contrast to the conventional procedure of covering the whole band with a single radiation mode. This proposal was a simple low-cost mechanism to improve connectivity significantly when the link is dynamic. Using the proposed method, the performance of this antenna was evaluated and its benefits were quantified over two conventional widely used antenna types with single radiation mode operation, a directional patch and an omnidirectional monopole.

To verify that the modelling process, the real world wireless performance of the proposed antenna was experimentally evaluated in a domestic environment. In the experiment, an on-body sensor node communicated with a nearby AP in the 2.4 GHz band using BLE.

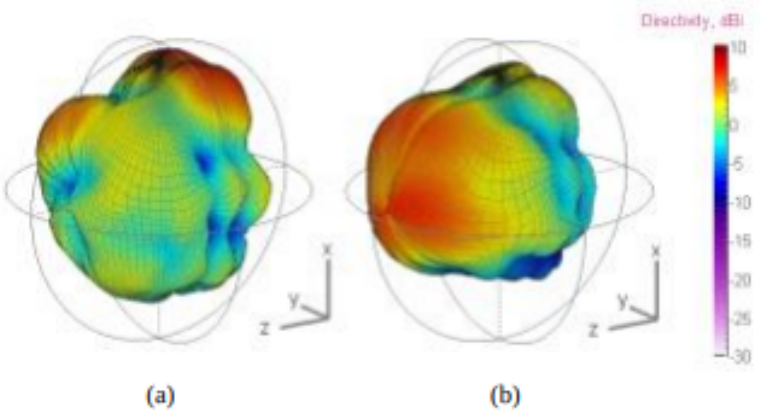


Figure 3.4: 3D Directivity (dBi) plots of (a) the azimuth mode and (b) the elevation mode.

3.2.2 Proposed Method

Human bodies are mobile which makes body-centric networks dynamic. Since the design of an antenna will be constant, optimisations for performance where a person is in a particular room, in a particular pose, might negatively affect antenna performance where the person is somewhere else doing something else. This is why existing models are inappropriate for generating meaningful performance metrics. The antenna evaluation procedure advocated in this work is shown in Fig. 3.5. This is a holistic method for developing antenna designs which considers the antenna performance in a dynamic multipath environment, rather than in isolation.

3.2.3 A Novel Antenna Design

A new type of antenna that is able to switch between radiation modes was designed. Details may be found in [14]. To produce a suitable antenna radiation model, accurate measurements of the antenna pattern, including the effects of the body were required. The antenna measurement campaign conducted for this work included anechoic chamber measurements of the antenna attached to a phantom in various positions. The radiation pattern is shown for the two modes in Fig 3.4. It can be observed that one mode has peak directivity in the azimuth direction and the other mode has peak directivity in the z direction.

The use of rotating turntables allowed a three-dimensional antenna pattern to be generated that includes the effects of being near the body. Computer scripts were used to rotate antenna patterns, with two polarisations, in 3D space to any required orientation.

3.2.4 Ray Tracer

The ray tracer that was used included reflection, refraction and diffraction of rays and had been verified against measurements in [35][36][32]. Used at the frequencies of interest, the ray-tracer

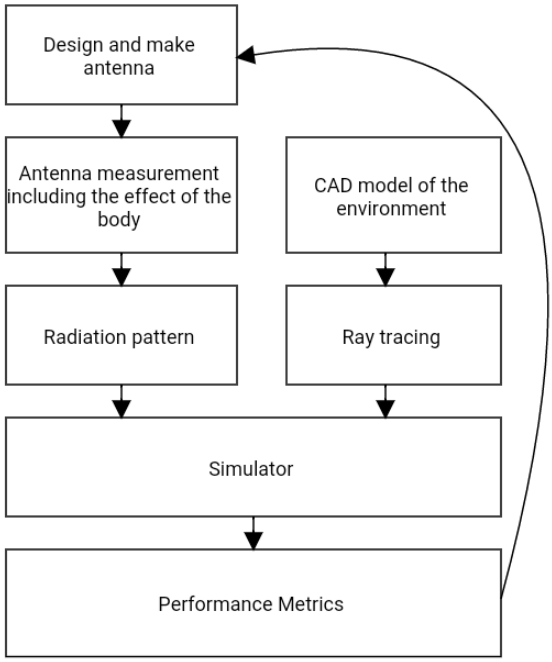


Figure 3.5: An overview of the proposed method for antenna design and verification.

could be considered state-of-the-art. The ray-tracer took as an input a computer aided design (CAD) model of the environment of interest. For the purposes of this work, a model of a three-storey house on Woodland Road, Bristol, UK was used. This is shown in Fig 3.6 . The model included both the spatial layout of the building, as well as material constitutive parameters. Client and access point locations were specified as needed. Complex phasors representing the signal in two orthogonal polarisations were traced through the environment in the tool, being modified as necessary by interactions with building materials and by propagation through space.

Analysing the time of arrival and state of phasors at the receive antenna allowed for a power delay profile and other information, such as phase, polarisation and received signal strength to be derived. In ray-tracing, all antennas were initially assumed to be isotropic and of unity gain. Using angle of arrival and departure information the radiation characteristics obtained from chamber measurements were convolved with data generated from ray-tracing to simulate the whole channel. Care was taken to ensure that inaccuracies introduced into the model, for example, due to unrepresentative material parameters, were kept to a minimum. Whilst there will inevitably be some inaccuracies in any simulation, there are several benefits to the use of ray tracing simulations. Results are more accurate/applicable than statistical models since more information about the environment is included. There is no requirement for engineers to disrupt a space or perform measurements, which might require expensive equipment. Data collection can also be automated and

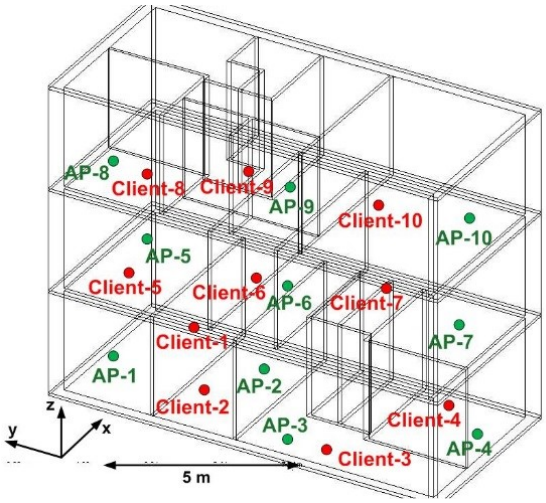


Figure 3.6: The test environment.

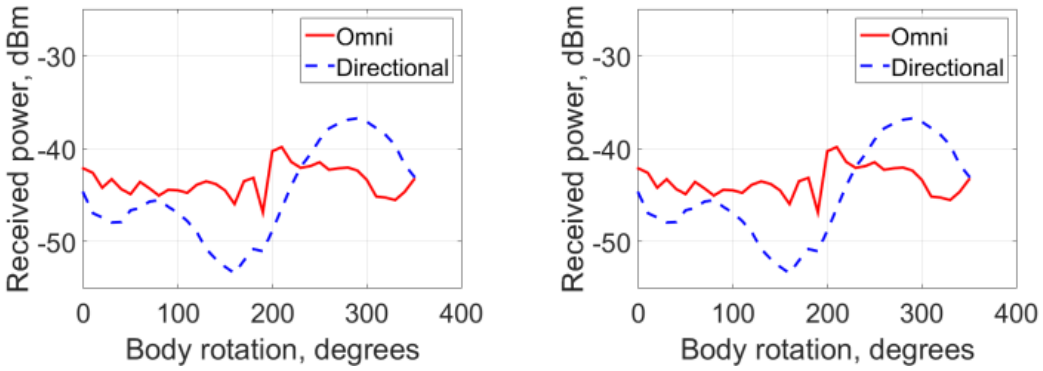


Figure 3.7: Received power at the user as a function of body rotation.(a) access-point and user in same room; (b) access-point and user in adjacent rooms.

results can be generated in any environment where it is possible to generate a CAD model. Data visualisations such as ray paths can also be inspected to identify significant propagation routes or particularly problematic links.

3.2.5 Simulated Antenna Performance Evaluation

The ray-tracing study was performed in the virtual three-storey test house shown in Fig. 3.6. The performance of the proposed antenna was compared to that of the single mode directional patch and a single mode omnidirectional monopole to quantify the benefits of the switching radiation mode technique. Results from ray-tracing were combined with the radiation models from chamber measurements to simulate the channel.

Initially, a ray-tracing algorithm predicted all the direct, reflected and diffracted rays at 2.44 GHz for two scenarios: the AP and the user are located in the same room (AP 1 – Client 1 in Fig.

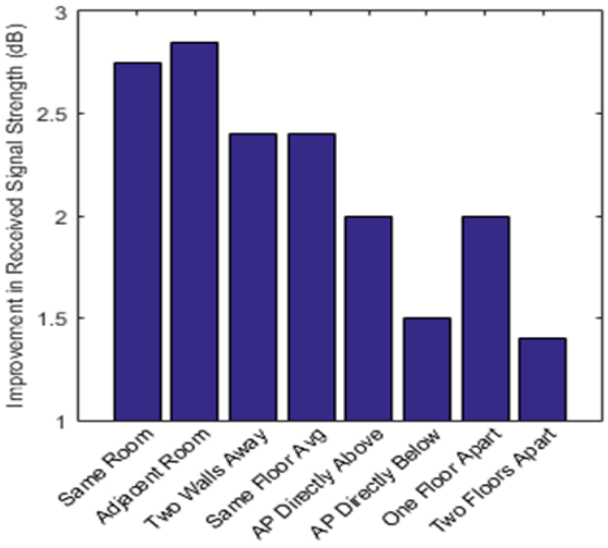


Figure 3.8: The proposed method.

3.6); the access-point and the user are located in adjacent rooms (AP 1 – Client 2 in Fig. 3.6). In both cases, the user was rotated through 360 degrees in azimuth, and the measured 3D polarimetric radiation patterns for a wrist-mounted omnidirectional and directional patch antenna were spatially convolved with the predicted multipath components. It should be noted that both elements had a similar measured radiation efficiency with that of the proposed antenna. The measured pattern of a vertical dipole was used at the side of the AP. Fig. 3.7 shows the received power at the user as a function of the body rotation for the two different antenna types in the two scenarios. When the user faced towards the direction of the dominant propagation path, the received signal strength was stronger by up to approximately 8 dB with the directional antenna, whereas, as the user turns towards other directions, the omnidirectional antenna provided more stable performance, better by up to about 8 dB than the directional antenna.

To obtain further insight into the received signal strength advantages of frequency dependent pattern switching, the ray-tracing simulation was used to analyse antenna performance in different topologies in the test house. Fig. 3.8 shows the improvement in received signal strength for different topologies achieved by switching between radiation modes when advantageous. Improvement is relative to the best single radiation mode on average for that topology. There was always an advantage to being able to switch between patterns regardless of topology with a maximum improvement of 2.9 dB. Greatest gains were achieved when the AP was nearer to the UE, with both separation by walls on the same floor, and separation by multiple floors decreasing the advantages of pattern switching. In the situation where AP and user were separated by two floors, the improvement in received signal strength due to pattern switching was only 1.4 dB.

A total of nine antenna positions were used for ray-tracing simulations. These were generated

by having the arm straight down by the side of the body (positions 1,4,7), a 45 degree bend at the elbow (positions 2,5,8) and a 90 degree bend at the elbow (positions 3,6,9). At each of these arm positions the antenna was mounted on the thumb side of the wrist, and the wrist was rotated clockwise 0 degrees (positions 7,8,9), 90 degrees (positions 4,5,6) or 180 degrees (positions 1,2,3). Fig. 3.9 shows these positions. Fig. 3.10 shows the difference in received signal strength between best and worst radiation mode. It should be noted there is no switching, the best radiation pattern on average relative to the worst is shown. Fig. 3.10 shows all antenna positions either with the AP on the same floor as the user or with the AP upstairs with respect to the user. In Fig. 3.10, if a bar is below zero on the dB scale it represents the directional pattern being better, in terms of received signal strength, than the omnidirectional pattern, and if above it represents the omnidirectional pattern's improvement over the directional pattern. Taking the example situation where the AP is upstairs relative to the user if the antenna is in position 2, an omnidirectional radiation pattern results in an extra 5.9 dB of signal strength compared to the directional pattern. If a user were to move their arm such that the antenna was in position 9, now the directional pattern achieves a 4.7 dB improvement in received signal strength over the omnidirectional. Another example is if the antenna is in position 4 on the same floor as the AP there is a 7.3 dB advantage to using the directional pattern, however if the user were to move downstairs relative to the AP with the same antenna position, 4, there is a 3.3 dB improvement from using the omnidirectional pattern. It can be observed that both arm position and topology significantly affects which radiation pattern is optimal. This motivates the use of the proposed antenna, which may switch between radiation modes in a normal domestic case in which users are mobile.

To simulate the proposed antenna design, an investigation was carried out to demonstrate the advantage of switching compared to picking a single radiation mode. By choosing the best radiation mode in a given situation the pattern switching antenna could be simulated. Fig. 3.11 shows how, in all elbow positions, pattern diversity facilitates an improved received signal strength versus both the worst and best case single pattern mode. In our case, this shows not only that there is an advantage to choosing a single best radiation mode, but that there is further significant advantage to switching between modes as and when this maximises received signal strength. A minimum improvement of 1.6 dB can be observed in Fig. 3.11 which is relative to a well-chosen single pattern antenna for a zero-degree elbow bend. A maximum of improvement of 4.3 dB is achieved compared to a poorly chosen single radiation pattern in the same elbow orientation.

Finally, taking all the simulations into account, the performance of the omnidirectional, the directional and the reconfigurable antenna proposed in this paper was compared for ten access-point and ten user locations randomly distributed in all the rooms around the test-house, as shown in Fig 3.6. In all cases, the user was rotated through 360 degrees in steps of 10 degrees and nine different wrist positions were considered. All analysis was repeated for three orthogonal dipoles at the access-point, resulting in a total of 94,500 links for each user antenna. This number of links

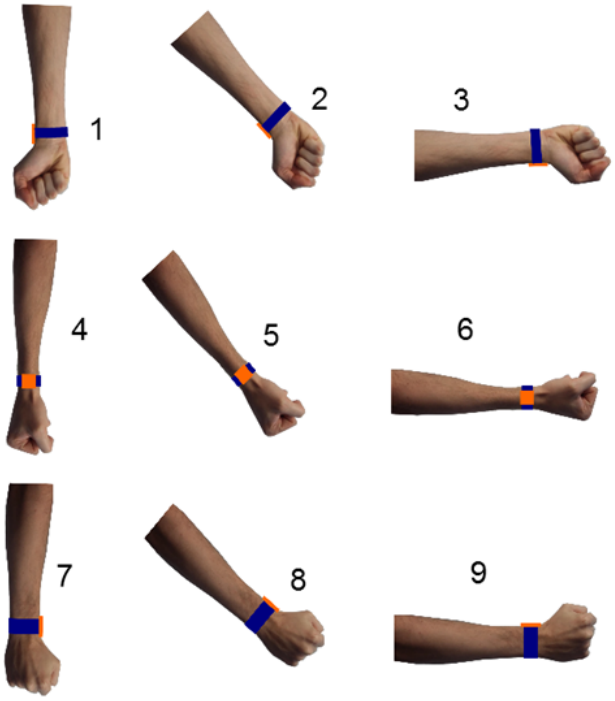


Figure 3.9: Antenna positions. The orange patch represents antenna position.

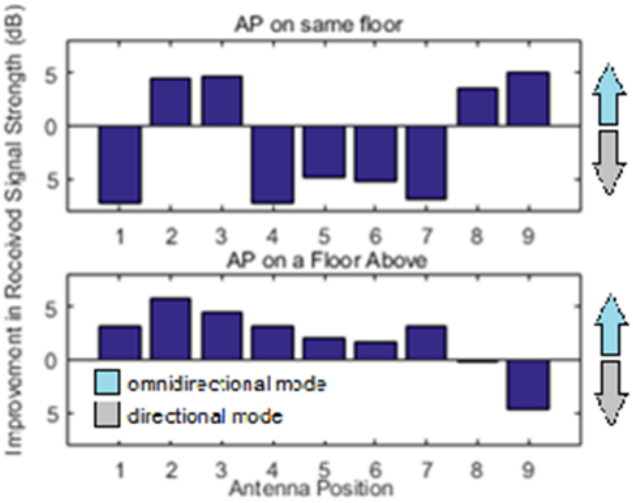


Figure 3.10: Difference in received signal strength due to choosing the radiation mode that gives the best performance on average for various antenna positions. Two topologies are shown. The figure demonstrates that an antenna design that could switch between radiation modes is desirable since different arm positions and topology cause optimal radiation mode to change.

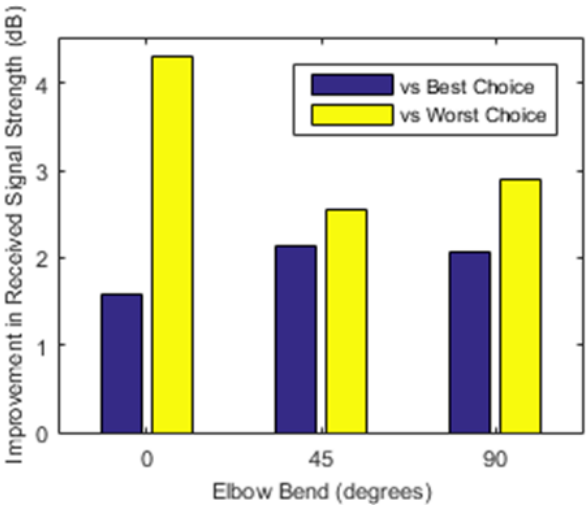


Figure 3.11: Improvement in received signal strength due to switching between radiation modes when advantageous relative to the best and worst single radiation modes for three elbow positions. This illustrates that the use of pattern diversity always results in an improved received signal strength when compared to an antenna with a single radiation mode.

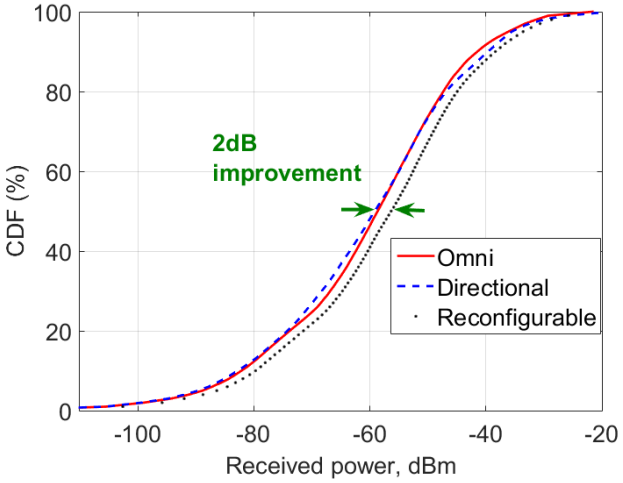


Figure 3.12: CDF of received power for the three different user antennas over the 94,500 links in the virtual test-house.

would be impractical to measure in the real world. Fig. 3.12 shows the Cumulative Distribution Function (CDF) of the received power for the three user antennas under test. When such a large number of links is investigated, differences tend to average out and hence and it can be seen that the CDF graphs for the omnidirectional and for the directional antennas are almost identical. The proposed antenna, however, shows a significant median improvement of approximately 2 dB due to frequency dependent pattern switching. This is a useful result for the system designer. The simulations have demonstrated the advantages of the example antenna. If results had shown poor performance compared to a single radiation mode the antenna designer could redesign the antenna until the specified performance was achieved. Useful insights into antenna performance in different situations are also gained. For example, omnidirectional antennas tend to give superior performance when communication takes place between floors, to a maximum of a 5 dB improvement, whereas directional antennas give better performance on the same floor unless the arm position is such that the main lobe is directed towards the floors above or below.

3.2.6 Verification Against Measurements

Ray tracing simulations were compared to real world measurements for verification purposes. Detailed analysis is omitted here, details may be found in [14].

3.2.7 Conclusions

In this chapter, a method for quantifying the performance of an antenna design was presented. As a case study, a planar antenna capable of radiation pattern switching by changing the maximum radiation direction with frequency was presented.

The performance of the proposed antenna was analysed using the proposed method which included extensive ray-tracing simulations in conjunction with antenna pattern measurements including the effects of the body. Early in the chapter, various ray tracing tools were presented. For the new method, a state of the art ray tracing tool was used illustrating the utility of such tools. A detailed investigation was performed to practically demonstrate the sensitivity of the link quality on the antenna position on the wrist and on the arm orientation. It was shown, that the frequency dependent pattern switching technique may result in large benefits, as high as 8 dB for specific links. The analysis of the antenna showed, in this test case, that the example antenna achieved good performance but could also be used to identify poor antenna designs or compare candidate antennas in specific situations.

Then the performance of the antenna was investigated in a real off-body propagation scenario in a residential environment. The proposed antenna was mounted on the wrist of a human subject and communicated with a nearby access-point. The impact of body-shadowing was quantified, showing that the standard deviation of the received signal due to body rotation is approximately

6 dB in links where a strong signal path is present, and about 4 dB in links where weaker multipaths are the main propagation mechanism. Furthermore, the benefit of pattern switching was highlighted as the median performance of the proposed antenna was found to be 2 dB better than that of a patch antenna with directional pattern only, and that of a monopole antenna with omnidirectional pattern only over the whole band. This result was calculated over all the measured links in the house, access-point antennas, arm positions, and body orientations and was in agreement with the conclusions from the ray-tracing simulations; demonstrating how the ray-tracing simulations can be used to effectively quantify antenna performance in the design phase of an off-body communication system.

An intuitive idea, namely that an antenna capable of switching patterns would outperform an antenna only able to produce one radiation pattern was validated. Whilst the ray tracing results cannot be used directly to design an antenna, they can be used to quantify performance and identify means of increasing performance. For example, by identifying the typical directions of arrival of multipath components a designer could use an antenna design known have high directivity in those directions. The controlled test environment offered by simulations allows any new design to be regression tested in scenarios where it was known that a previous design performed well and can be evaluated in previously problematic scenarios. Because simulations are cheap in terms of effort, very large scale campaigns can be run generating statistically robust data mimicking real world use.

Bibliography

[1] D. Gallagher (2008). photonics CAD matures - Google search. `https://www.google.co.uk/webhp?sourceid=chrome-instant&ion=1&espv=2&ie=UTF-8#safe=off&q=D.+Gallagher+(2008).+Photonics+CAD+Matures`. (Accessed on 07/07/2016).

[2] Wireless EM propagation software - Wireless Insite.

[3] F Aguado Agelet, Arno Formella, JM Hernando Rabanos, F Isasi De Vicente, and F Perez Fontan. Efficient ray-tracing acceleration techniques for radio propagation modeling. *IEEE transactions on Vehicular Technology*, 49(6):2089–2104, 2000.

[4] Zahid Aslam, Yoann Corre, Emil Björnson, and Erik G. Larsson. Performance of a dense urban massive MIMO network from a simulated ray-based channel. *EURASIP Journal on Wireless Communications and Networking*, 2019, 12 2019.

[5] Lee BS. *Thesis*. PhD thesis, University of Bristol, 2002.

[6] Anupam R. Chandran, Gareth A. Conway, and William G. Scanlon. Pattern switching compact patch antenna for on-body and off-body communications at 2.45 GHz. In *2009 3rd European Conference on Antennas and Propagation*, European Conference on Antennas and Propagation. Proceedings, pages 2055–2057. IEEE, 3 2009.

[7] D. J. Cichon, T. C. Becker, and M. Dottling. Ray optical prediction of outdoor and indoor coverage in urban macro- and micro-cells. In *Vehicular Technology Conference, 1996. Mobile Technology for the Human Race., IEEE 46th*, volume 1, pages 41–45 vol.1, Apr 1996.

[8] Y. Corre and Y. Lostanlen. Three-dimensional urban EM wave propagation model for radio network planning and optimization over large areas. *IEEE Transactions on Vehicular Technology*, 58(7):3112–3123, Sep. 2009.

[9] I. J. Craddock, D. L. Paul, C. L. Railton, P. N. Fletcher, and M. Dean. Applications of single mode extraction from finite difference time domain data. *IEE Proceedings - Microwaves, Antennas and Propagation*, 146(2):160–162, Apr 1999.

[10] V. Degli-Esposti, F. Fuschini, E. M. Vitucci, M. Barbiroli, M. Zoli, L. Tian, X. Yin, D. A. Dupleich, R. Müller, C. Schneider, and R. S. Thomä. Ray-tracing-based mm-wave beamforming assessment. *IEEE Access*, 2:1314–1325, 2014.

[11] V. Degli-Esposti, F. Fuschini, E. M. Vitucci, and G. Falciasecca. Measurement and modelling of scattering from buildings. *IEEE Transactions on Antennas and Propagation*, 55(1):143–153, Jan 2007.

[12] V. Degli-Esposti, D. Guiducci, A. de'Marsi, P. Azzi, and F. Fuschini. An advanced field prediction model including diffuse scattering. *IEEE Transactions on Antennas and Propagation*, 52(7):1717–1728, July 2004.

[13] U. Dersch and E. Zollinger. Propagation mechanisms in microcell and indoor environments. *IEEE Transactions on Vehicular Technology*, 43(4):1058–1066, Nov 1994.

[14] S. Dumanli, L. Sayer, E. Mellios, X. Fafoutis, G. S. Hilton, and I. J. Craddock. Off-body antenna wireless performance evaluation in a residential environment. *IEEE Transactions on Antennas and Propagation*, 65(11):6076–6084, Nov 2017.

[15] Mellios Evangelos. *Physical Layer Analysis of Communication Systems Integrating Antenna Characteristics and Propagation Modelling for Indoor and Outdoor Environments*. PhD thesis, University of Bristol, 2013.

[16] Franco Fuschini, M Zoli, Enrico Vitucci, M Barbiroli, and Vittorio Degli-Esposti. A study on mm-wave multi-user directional beamforming based on measurements and ray tracing simulations. *IEEE Transactions on Antennas and Propagation*, PP:1–1, 01 2019.

[17] P. S. Hall, Y. Hao, Y. I. Nechayev, A. Alomainy, C. C. Constantinou, C. Parini, M. R. Kamarudin, T. Z. Salim, D. T. M. Hee, R. Dubrovka, A. S. Owadally, W. Song, A. Serra, P. Nepa, M. Gallo, and M. Bozzetti. Antennas and propagation for on-body communication systems. *IEEE Antennas and Propagation Magazine*, 49(3):41–58, June 2007.

[18] M. Hata. Empirical formula for propagation loss in land mobile radio services. *IEEE Transactions on Vehicular Technology*, 29(3):317–325, Aug 1980.

[19] C. M. P. Ho and T. S. Rappaport. Wireless channel prediction in a modern office building using an image-based ray tracing method. In *Global Telecommunications Conference, 1993, including a Communications Theory Mini-Conference. Technical Program Conference Record, IEEE in Houston. GLOBECOM '93., IEEE*, pages 1247–1251 vol.2, Nov 1993.

[20] T. Holt, K. Pahlavan, and J. F. Lee. A graphical indoor radio channel simulator using 2D ray tracing. In *Personal, Indoor and Mobile Radio Communications, 1992. Proceedings, PIMRC '92., Third IEEE International Symposium on*, pages 411–416, Oct 1992.

[21] W. Honcharenko, H. L. Bertoni, J. L. Dailing, J. Qian, and H. D. Yee. Mechanisms governing UHF propagation on single floors in modern office buildings. *IEEE Transactions on Vehicular Technology*, 41(4):496–504, Nov 1992.

[22] R. Hoppe, G. Wölfle, and U. Jakobus. Wave propagation and radio network planning software WinProp added to the electromagnetic solver package FEKO. In *2017 International Applied Computational Electromagnetics Society Symposium - Italy (ACES)*, pages 1–2, March 2017.

[23] S. Hussain and C. Brennan. An efficient ray tracing method for propagation prediction along a mobile route in urban environments. *Radio Science*, 52(7):862–873.

[24] S. Hussain and C. Brennan. An efficient ray-tracing acceleration technique for mobile receivers in urban environments. In *2017 Computing and Electromagnetics International Workshop (CEM)*, pages 47–48, June 2017.

[25] F. Ikegami, T. Takeuchi, and S. Yoshida. Theoretical prediction of mean field strength for urban mobile radio. *IEEE Transactions on Antennas and Propagation*, 39(3):299–302, Mar 1991.

[26] MF Iskander, Zhengqing Yun, and Zhijun Zhang. Outdoor/indoor propagation modeling for wireless communications systems. In *Antennas and Propagation Society International Symposium, 2001. IEEE*, volume 2, pages 150–153. IEEE, 2001.

[27] itu. *ITU-R Recommendation*.

[28] S. Kang and C. W. Jung. Wearable fabric antenna on upper arm for medradio band applications with reconfigurable beam capability. *Electronics Letters*, 51(17):1314–1316, 2015.

[29] M. C. Lawton, R. L. Davies, and J. P. McGeehan. A ray launching method for the prediction of indoor radio channel characteristics. In *Personal, Indoor and Mobile Radio Communications., IEEE International Symposium on*, pages 104–108, Sep 1991.

[30] M. C. Lawton and J. P. McGeehan. The application of GTD and ray launching techniques to channel modelling for cordless radio systems. In *Vehicular Technology Conference, 1992, IEEE 42nd*, pages 125–130 vol.1, May 1992.

[31] M. C. Lawton and J. P. McGeehan. The application of a deterministic ray launching algorithm for the prediction of radio channel characteristics in small-cell environments. *IEEE Transactions on Vehicular Technology*, 43(4):955–969, Nov 1994.

[32] B. S. Lee, A. R. Nix, and J. P. McGeehan. Indoor space-time propagation modelling using a ray launching technique. In *Antennas and Propagation, 2001. Eleventh International Conference on (IEE Conf. Publ. No. 480)*, volume 1, pages 279–283 vol.1, 2001.

[33] B. S. Lee, A. R. Nix, and J. P. McGeehan. A spatio-temporal ray launching propagation model for UMTS pico and microcellular environments. In *Vehicular Technology Conference, 2001. VTC 2001 Spring. IEEE VTS 53rd*, volume 1, pages 367–371 vol.1, 2001.

[34] B. S. Lee, C. M. Tan, S. E. Foo, A. R. Nix, and J. P. McGeehan. Site specific prediction and measurement of indoor power delay and power azimuth spectra at 5 GHz. In *Vehicular Technology Conference, 2001. VTC 2001 Fall. IEEE VTS 54th*, volume 2, pages 733–737 vol.2, 2001.

[35] B. S. Lee, C. M. Tan, S. E. Foo, A. R. Nix, and J. P. McGeehan. Site specific prediction and measurement of indoor power delay and power azimuth spectra at 5 GHz. In *IEEE 54th Vehicular Technology Conference. VTC Fall 2001. Proceedings (Cat. No.01CH37211)*, volume 2, pages 733–737 vol.2, Oct 2001.

[36] BS Lee, AR Nix, and JP McGeehan. A spatio-temporal ray launching propagation model for UMTS pico and microcellular environments. In *IEEE VTC Spring 2001, Rhodes, May*, volume 1, pages 367 – 371, United States, 5 2001. Institute of Electrical and Electronics Engineers (IEEE).

[37] Z. G. Liu and Y. X. Guo. Dual band low profile antenna for body centric communications. *IEEE Transactions on Antennas and Propagation*, 61(4):2282–2285, April 2013.

[38] G. R. Maccartney, T. S. Rappaport, S. Sun, and S. Deng. Indoor office wideband millimeter-wave propagation measurements and channel models at 28 and 73 ghz for ultra-dense 5g wireless networks. *IEEE Access*, 3:2388–2424, 2015.

[39] J.W. McKown and R.L. Hamilton. Ray tracing as a design tool for radio networks. In *IEEE Network Magazine, Vol. 5, No. 6, Nov. 1991*.

[40] Aleksandar Neskovic, Natasa Neskovic, and George Paunovic. Modern approaches in modeling of mobile radio systems propagation environment. *IEEE Communications Surveys and Tutorials*, 3(3):2–12, 2000.

[41] K. H. Ng, E. K. Tameh, A. Doufexi, M. Hunukumbure, and A. R. Nix. Efficient multielement ray tracing with site-specific comparisons using measured mimo channel data. *IEEE Transactions on Vehicular Technology*, 56(3):1019–1032, May 2007.

[42] J. P. Rossi, J. C. Bic, A. J. Levy, Y. Gabillett, and M. Rosen. A ray launching method for radio-mobile propagation in urban area. In *Antennas and Propagation Society International Symposium, 1991. AP-S. Digest*, pages 1540–1543 vol.3, June 1991.

[43] K. R. Schaubach, N. J. Davis, and T. S. Rappaport. A ray tracing method for predicting path loss and delay spread in microcellular environments. In *Vehicular Technology Conference, 1992, IEEE 42nd*, pages 932–935 vol.2, May 1992.

[44] J. W. Schuster and R. J. Luebbers. Comparison of GTD and FDTD predictions for UHF radio wave propagation in a simple outdoor urban environment. In *Antennas and Propagation Society International Symposium, 1997. IEEE., 1997 Digest*, volume 3, pages 2022–2025 vol.3, July 1997.

[45] S. Y. Seidel and T. S. Rappaport. A ray tracing technique to predict path loss and delay spread inside buildings. In *Global Telecommunications Conference, 1992. Conference Record., GLOBECOM '92. Communication for Global Users., IEEE*, pages 649–653 vol.2, Dec 1992.

[46] Scott Y Seidel and Theodore S Rappaport. Site-specific propagation prediction for wireless in-building personal communication system design. *IEEE transactions on Vehicular Technology*, 43(4):879–891, 1994.

[47] Q. H. Spencer, B. D. Jeffs, M. A. Jensen, and A. L. Swindlehurst. Modeling the statistical time and angle of arrival characteristics of an indoor multipath channel. *IEEE Journal on Selected Areas in Communications*, 18(3):347–360, March 2000.

[48] Julien Stephan, Yoann Corre, Romain Charbonnier, and Yves Lostanlen. Increased reliability of outdoor millimeter-wave link simulations by leveraging lidar point cloud. 04 2018.

[49] W. K. Tam and V. N. Tran. Multi-ray propagation model for indoor wireless communications. *Electronics Letters*, 32(2):135–137, Jan 1996.

[50] E. K. Tameh and A. R. Nix. The use of measurement data to analyse the performance of rooftop diffraction and foliage loss algorithms in a 3-d integrated urban/rural propagation model. In *VTC '98. 48th IEEE Vehicular Technology Conference. Pathway to Global Wireless Revolution (Cat. No.98CH36151)*, volume 1, pages 303–307 vol.1, May 1998.

[51] J. H. Tarng and T. R. Liu. Effective models in evaluating radio coverage on single floors of multifloor buildings. *IEEE Transactions on Vehicular Technology*, 48(3):782–789, May 1999.

[52] R. P. Torres, L. Valle, M. Domingo, S. Loredo, and M. C. Diez. Cindoor: an engineering tool for planning and design of wireless systems in enclosed spaces. *IEEE Antennas and Propagation Magazine*, 41(4):11–22, Aug 1999.

[53] R. Valenzuela. A ray tracing approach to predicting indoor wireless transmission. In *Vehicular Technology Conference, 1993., 43rd IEEE*, pages 214–218, May 1993.

[54] Ying Wang, Safieddin Safavi-Naeini, and Sujeet K Chaudhuri. A hybrid technique based on combining ray tracing and fdtd methods for site-specific modeling of indoor radio wave propagation. *IEEE Transactions on antennas and propagation*, 48(5):743–754, 2000.

[55] Kane Yee. Numerical solution of initial boundary value problems involving maxwell's equations in isotropic media. *IEEE Transactions on Antennas and Propagation*, 14(3):302–307, May 1966.

[56] L. Zhang, Z. Wang, and J. L. Volakis. Textile antennas and sensors for body-worn applications. *IEEE Antennas and Wireless Propagation Letters*, 11:1690–1693, 2012.

[57] Zhijun Zhang, R. K. Sorensen, Zhengqing Yun, M. F. Iskander, and J. F. Harvey. A ray-tracing approach for indoor/outdoor propagation through window structures. *IEEE Transactions on Antennas and Propagation*, 50(5):742–748, May 2002.

Chapter 4

Ray Launching Tool

Some results in this chapter were obtained in collaboration with Nick Burris of Shanghai University. Specifically, section 4.6.2 contains collaborative work.

4.1 Motivation

In previous chapters, it was explained that new propagation models are needed because, contrary to previous generations of cellular radio, 5G networks will be inherently 'spatial'. It was also argued that ray modelling is suitable for this task, and that the use of ray modelling results can augment the design of devices. Several competing ray tools, commercial or academic were introduced in the previous chapter. The problem with many of these tools was that they were either proprietary, precluding the possibility of further customisation, or too mature to change without significant refactoring. A new ray launching tool was designed which provided a platform for further research, including experimental modelling functions. A description of the design of the tool is provided in this chapter.

4.2 Overview

A system level diagram of the developed ray launching software is given in Fig 4.1.The various input and output files are shown as well as indicating some of the internal processes in the software.

4.2.1 Inputs

Referring to Fig. 4.1, the software was designed to take several input files.

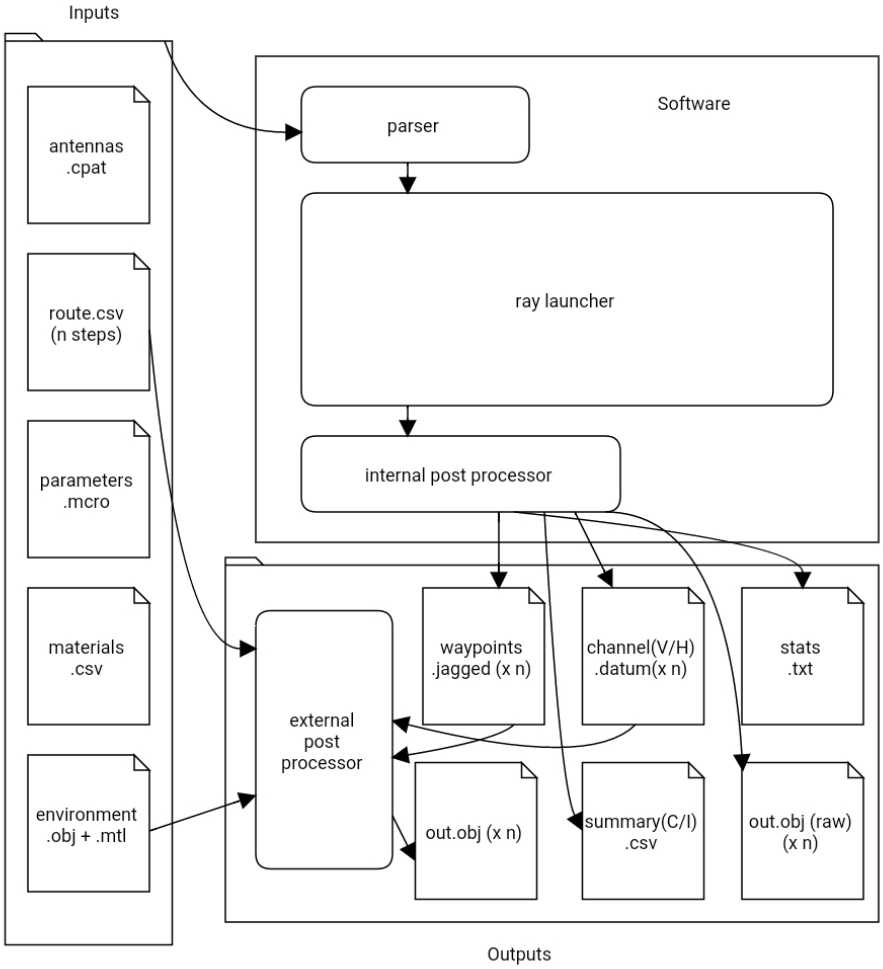


Figure 4.1: System overview of the developed tool set including inputs and outputs.

1. Two .cpat files. Each file specified, for a particular polarisation, the radiation pattern of an antenna. More information on the file format may be found in Chapter 3.

2. The route file. This file was comprised of one or more lines specifying a TX location and an RX location and their respective orientations and velocities.

3. A macro file specified various setup parameters. The various parameters are discussed separately later in this section.

4. The materials file specified material properties such as permeability, permittivity, roughness and velocity.

5. The environment .obj file was the model of the environment. The associated .mtl file contained the names and rendering instructions for the various materials associated with faces by tagging in the .obj file.

6. An intermediate material lookup file (pseudo input) allowed the values in the materials file to be associated with those in the .mtl file. This enabled the software to make use of standard 3D model formats whilst adding extra physical parameters.

The parameters macro file contained all the setup parameters for the software. The format of the file was text based and it followed a key and value structure. This meant that the order of the parameters in the file could be changed by a user with no deleterious effects on the software. The parameters included:

1. File paths to the other input files.

2. The frequency of interest.

3. Units of the .obj file (i.e. is '1' 1 m or 1 mm).

4. The number of rays to be fired.

5. The maximum number of interactions before a ray path was terminated.

6. Whether to reload the .obj file (parsing was time intensive, often one might have wished to reload data from processed lists of vertices and connections between them from a previous simulation).

7. The polarisation of the excitation of the transmitting antenna.

8. Tokens specifying which propagation phenomena should be simulated. If scattering was turned on, then a scattering model also had to be specified.

4.2.2 Outputs

For every line in the route file, a .jagged file and two .datum files were produced. These file formats were a custom text based format, chosen to be compatible with a legacy piece of ray tracing software [2]. Each .jagged file was as many lines long as ray paths that had been found for the particular part of the route corresponding to the file. Each line was a series of comma separated triplets specifying endpoints of the line segments forming a ray path. This information was purely geometric. The .datum files contained more quantitative information about both geometry and electromagnetic contributions of the ray path. Total distance, total time, free-space-path-loss over that distance, loss including antenna gains and interactions with the environment, phase, angles of arrival and departure and Doppler shifts were all specified in a .datum file. There was one .datum file for horizontal polarisation and one for vertical polarisation. Lines in the .datum files corresponded to lines in the .jagged file so the data in each could be easily combined in post processing if necessary. A statistics file was also produced. For each step on the route the statistics file contained various statistics like angular spreads, delay spread, coherence time, Doppler spread etc. Two summary files had as many lines as the route file. Once specified a coherent (vector) sum of rays for each step, the other an intensity sum.

In Fig 4.1, two types of output .obj files are shown. Both types were copies of the input .obj file but with ray paths added following the simulations. A new output file was produced for each step on the route, so if a multi-step route was simulated there would be a set of .obj files. The raw output .obj files just showed all found ray paths. This was sometimes useful for debugging, but actually was not useful for determining which paths were actually significant. An external post processing application was designed to filter ray paths so that only ray paths within some specified range from the most powerful path were shown in a second set of .obj files. In this processed set, ray paths were also colour coded, lighter green rays encoding more powerful ray paths and darker blue rays less powerful ray paths. Also shown in the output .obj files were antenna location markers. These were little colour coded axis markers, which helped to visualise the location and rotation of the antenna for verification purposes. These are shown in Fig 4.2 and Fig 4.3.

4.3 The Ray Launching Core

4.3.1 Ray Direction Initialisation

The tool was the pin cushion type of ray launcher. This meant that rays needed to be "fired" as evenly as possible from a transmitter location. The method usually found in literature for generating appropriate ray directions is the geometric subdivision method. When performing the method, the triangular faces of a geodesic structure are subdivided into more triangles. The new vertices produced are projected onto the unit sphere to get ray directions. This method is used in [1][4][5].

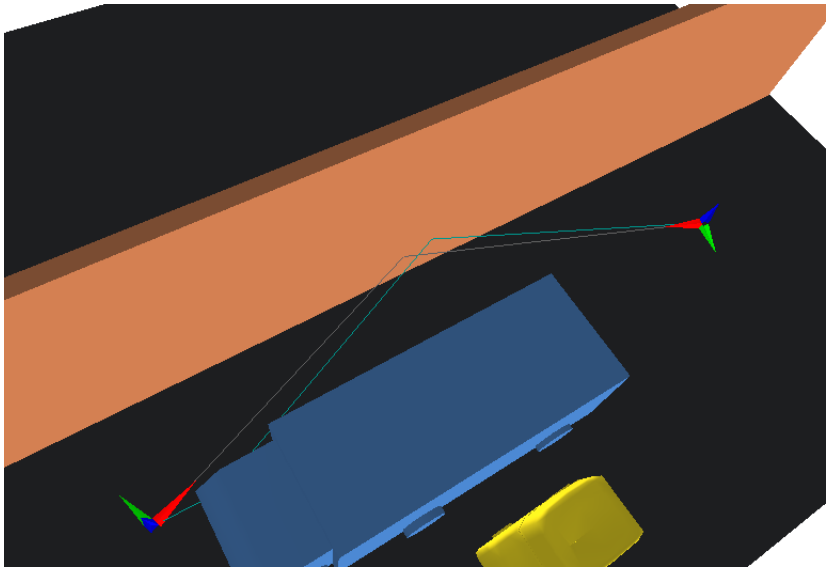


Figure 4.2: A post processed obj file, with two ray paths of different powers, and transceiver markers. The markers show the directions of an antenna referenced local coordinate system. This allows antenna rotations to be visualised and verified. The main beam of the antenna radiation pattern is aligned with the red marker.

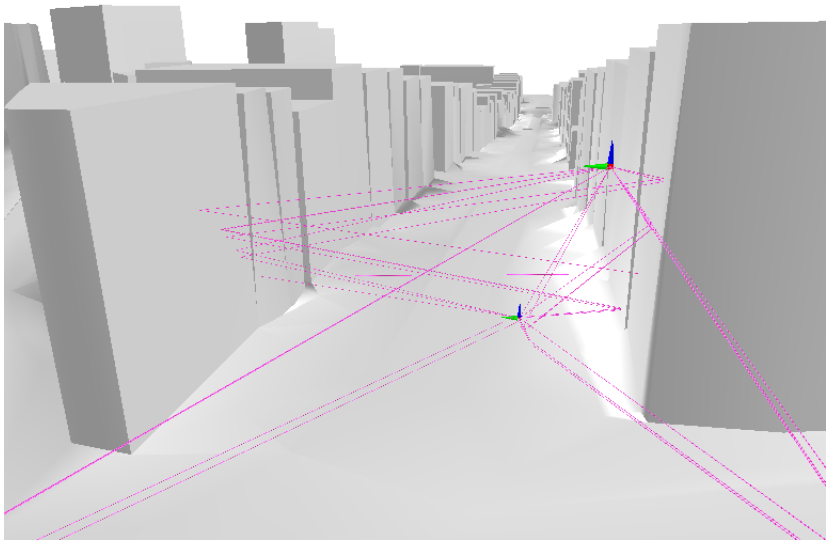


Figure 4.3: A raw .obj file showing all ray paths as equally significant. Diffuse scattering is not turned on, if it were, there would be a very high density of rays and this kind of plot would become difficult to view. The buildings shown are three stories, and approximately 9 m tall. The total area is in the order of 1 km by 1 km.

There are a few disadvantages to this method. The first disadvantage is that one is limited to certain discrete jumps in the number of rays. Every time the geodesic structure is subdivided, the number of rays goes up by a factor of three. Another disadvantage is that the distribution of points becomes uneven due to the subdivision process inherent to the method. Finally this method was found to be slow to compute in practice.

The tool uses an alternative method. The method is based on a "golden spiral". Given a number of rays, n_R, the matrix O, where O_i is row i,is calculated using the following algorithm.

Algorithm 1 Algorithm to return a list of initial ray direction vectors

$j \leftarrow \frac{\pi}{3-\sqrt{5}}$
$a \leftarrow 0$
$d \leftarrow \frac{2}{n_R}$
$z \leftarrow 1 - \frac{d}{2}$
for $i = 0$ to $i = n_R$ do
  $p \leftarrow \sqrt{1 - z^2}$
  $m \leftarrow \sqrt{z^2 + (p\cos(a))^2 + (p\sin(a))^2}$
  $O_i \leftarrow [p\cos(a), p\sin(a), z]/m$
  $z \leftarrow z - d$
  $a \leftarrow a + j$
end for
return O

Each row of O, O_i, is the direction vector for the i^{th} ray. The advantages of this method are threefold. It is quick to execute. Any number of rays may be fired. The distribution of points is more even.

4.3.2 Recursive Routine

In Fig 4.7, a flow diagram is shown which shows the basic control flow of the main ray launching function. The function was recursive, and took as an input an infant ray with a start point and direction. Some details have been omitted for clarity, for example, there is the possibility of turning off any propagation mechanism, and in Fig 4.7 it is assumed all are turned on.

The function first intersected the ray with the scene if the maximum interaction order had not yet been reached, else it terminated. If the ray did not hit the mesh, it may still have hit diffraction edges or the target before propagating forever into space. To handle such situations, a ray segment was formed by connecting the start point of the infant ray to some point a large distance away, in the direction of the ray. The segment from start point to point where it intersects the mesh - or to the point a very large distance away - was checked for prior interaction with the target or diffraction edge. If the target was hit first, a valid ray path was found and the procedure shown

Figure 4.4: Three hundred rays

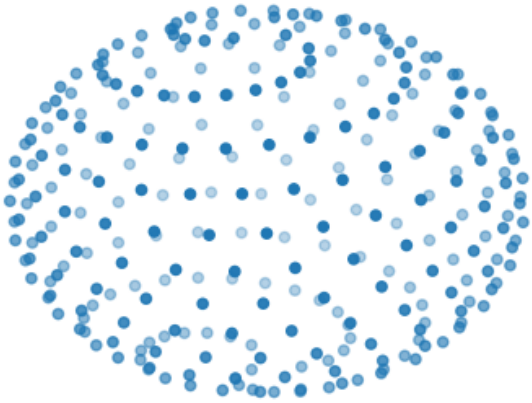


Figure 4.5: Nine hundred rays

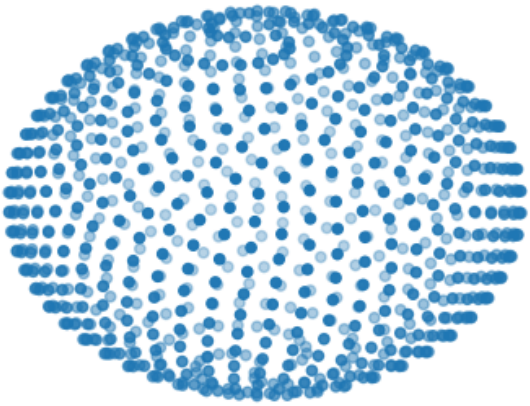
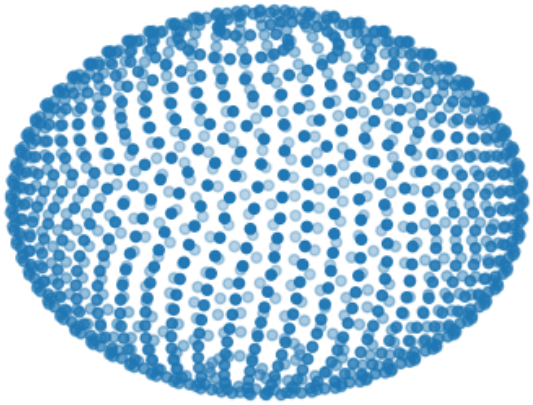


Figure 4.6: Twelve hundred rays



in Fig 4.14 was initiated and the routine terminated. This routine will be discussed in more detail later. If the target was not hit first, but a diffraction edge was, a loop was entered that generated all the necessary diffracted rays around the diffraction cone. Each time a new one was generated the recursive function was called on the new ray. Once all diffracted rays had been produced, or if an edge was not hit, the distance to the intersect point was checked. If it was not very large (i.e. the ray hasn't propagated out into space but has hit the mesh) a reflected ray, transmitted ray and scattered ray were all generated, with the recursive function called on the new rays as each was generated.

The method for generating diffracted rays is presented in the next chapter, reflected and transmitted rays were generated with a direction based on Snell's law. The direction of a scattered ray demands more attention. At each scattering element radiation is scattered in every direction. Firing rays in every direction from each scattering element would cause an exponentially increasing number of rays which would exhaust computational resources very quickly. Instead, the assumption was made that a scatter of a scatter would be unimportant. This is likely because energy respreads from each scattering element meaning propagation losses are multiplicative rather than additive. At each scatter element, one scattered ray was launched, which was always directed towards the transceiver point from which the rays had not emanated. If a token was set, indicating the last interaction was a scattered ray, no further interactions were simulated. This limited otherwise incredible computational complexity, where scatters of scatters of scatters would need to be accounted for. To maintain channel reciprocity, which would be violated when handling scattering this way otherwise, the whole function was repeated, but with the emanating and receiving point switched. Only paths that contained a scattering interaction - and with more than one interaction in total - were stored in this reverse operation. This meant that scatters may have occurred on the first interaction after leaving the TX and the last interaction before arriving at the RX.

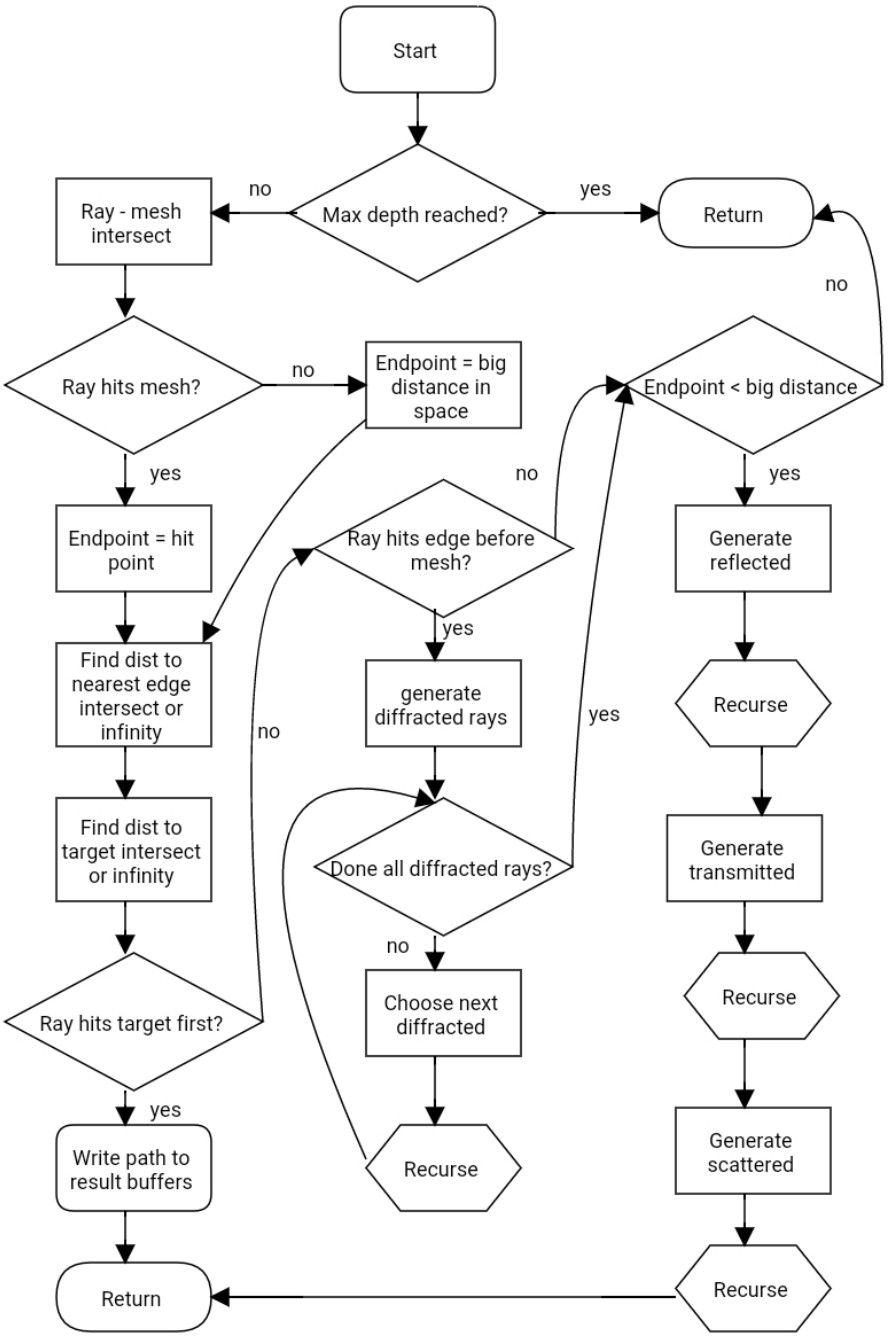


Figure 4.7: Programme control flow for the ray launching tool.

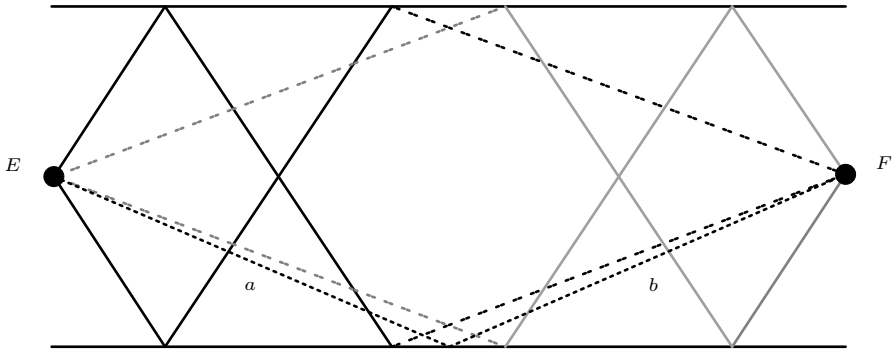


Figure 4.8: The forward and backward path finding process maintains channel reciprocity.

Fig 4.8 illustrates some of the ideas just described. In Fig 4.8, two transceiver locations are shown, E and F. Two of the paths shown start at E and propagate to F in one reflection and one scatter. These paths are shown in black, and for both paths there is a single reflection, surrounded by solid-line ray-segments, followed by a scatter, where the segment connecting the scattering point to F is shown as a dashed line. These paths would be found in the forward path finding operation. These paths have duals, indicated in light grey, where the scatter happens directly after leaving E, with dashed lines to the scattering point, and then there is a reflection, shown as solid lines before arriving at F. These duals are found by the backwards path finding step. The dotted path, labelled a and b is a single scatter interaction - which happens to be in the specular direction but need not be. It would be found in the forward path finding step. It also has a dual, which is just itself. It would be an error to find this path twice, so any single scatters are omitted from the backwards path finding step.

4.3.3 Discerning New Material

When a ray undergoes an interaction, it either stays in the material it was previously in, or propagates into a new one. For example, for a ray in free space impinging upon a brick wall, the reflected ray continues in free space, whilst the transmitted ray propagates into brick. Normally it is reasonably obvious what the new material should be based on the type of interaction. However, when modelling an environment comprised of solids, there is potential to make small "overlap" errors, for example a region of brick may slightly overlap a region of wood. Even if utmost care has been taken to ensure there is no overlapping, if two regions have co-planar faces, then it is likely that numerical errors due to finite precision on any realisable machine will cause a particular new material to be falsely identified, and hence ruin propagation predictions. The method for avoiding this type of error is described in this subsection.

The initial medium is stored. This will usually be air. Rays are fired into the environment and intersect with a triangle of some material. If in air, the situation is slightly different than if in a material of some other type, because the region of air is not usually explicitly modelled, so there are no faces made of air to intersect with. When a ray in air impinges upon a face, a "backray is fired". A "backray" is a ray that has opposite direction to the impinging ray that has its origin a small distance beyond the intersect point of the original ray and the triangle it intersects. In the case of a ray that was in air, this either hits the same triangle as the one previously intersected, in which case a transmitted wave is moving into a material of that triangles type, or it hits a different triangle of the same material, in which case there has been a very thin slice of material, and the ray is back out in air.

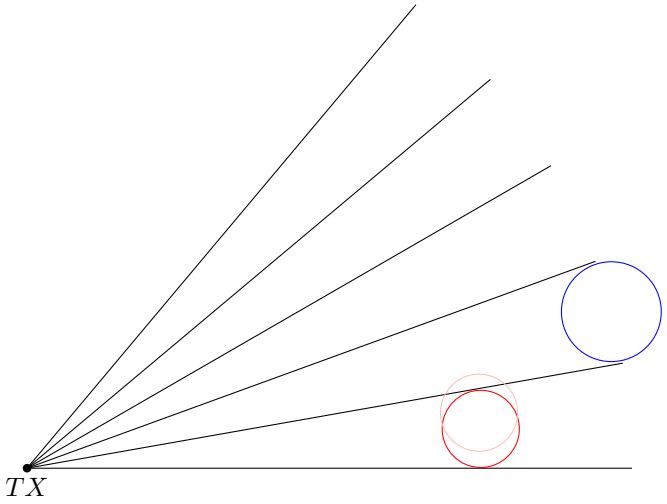
If in a material, if the backray hits the same triangle as hit by the original ray, this is a single skin boundary, which signifies the ray is moving back into air. If in some non-air material, and the backray hits a different triangle to the original ray, the situation needs to be discerned further. If both the triangles are of the same material, this signifies a boundary between two blocks of the same material, which is a common situation if a compound object has been constructed in modelling software. In this case, the ray stays in the same material. If it is found there are two different materials for the two different triangles, they are compared to the material the ray is currently in. The one that is not of the same type as the material the ray is currently in, is the material moved into.

4.3.4 Hitting the Target

In a ray launcher, rays are fired at a discrete angle. These rays, in general, do not intersect exactly an infinitesimal RX point. Rays are instead counted as "received" when they enter a region around the RX point, the size of which is dependent on the total distance travelled from the TX, and the angle between rays. This accounts for divergence. Fig 4.9 shows this dependence in two dimensions. At any distance from the TX the "reception circle", (or sphere in the three dimensional case) is just large enough that two rays, or three in three dimensions, separated by the angular resolution specified, intersect it. In Fig 4.9, there is a red solid circle which is just large enough to intersect two rays at a given distance from the TX. The solid pink circle represents the general case, where the circle is not exactly between two rays, but is of the same size as the red at the same distance. The solid blue circle is larger, because it is at a greater distance.

Spheres are used to decide if a ray is received, rather than a more complex shape based on the hexagonal cross section ray tubes, to simplify analysis and speed up intersection tests. This means some double counting of ray paths will occur. To make the sphere large enough to be skimmed by three rays, it must be large enough to receive two rays completely. Fig 4.10 illustrates the effect by showing how circles must overlap to fully cover a 2D plane. In three dimensions, the same effect is observed for spherical caps on a sphere. These overlapping sections, shown in darker red, cause

Figure 4.9: Reception circles



double counting of certain ray paths, and these double counted paths must be removed in post processing. In practice, this is achieved by checking rays for similarity based on several metrics, such as number of bounces, type of interactions, distance travelled and angle of arrival. Referring to Fig 4.11 with α the angle between rays, d the distance along a ray path. The distance p as a function of distance along the ray path, d is

$$p = d \sin\left(\frac{\alpha}{2}\right). \tag{4.1}$$

Using the assumption that the rays are equally spaced, and that enough rays are used that the curvature of the wavefront is negligible on a local scale, notice that the centroids of the ray tubes are separated by a mutually equal distance, so q can be determined.

$$q = \frac{2p}{\sqrt{3}}. \tag{4.2}$$

Next,

$$\beta = \arcsin(q). \tag{4.3}$$

Now referring to Fig 4.12, the relationship between q and the required radius, r, can be seen. To get the radius of the reception sphere, r,

$$r = d \tan(\beta). \tag{4.4}$$

Referring to Fig 4.13, the origin of the most recent ray segment is defined as $\hat{\mathrm{s}}$. Note that this need not be the TX point in a multi bounce path. Let the RX location be $\hat{\mathrm{e}}$, the ray direction $\hat{\mathrm{g}}$, and the path distance to $\hat{\mathrm{s}}$ be d_t and the variable distance from $\hat{\mathrm{s}}$ to the wavefront be d_n such that that $d = (d_t + d_n)$ then it can be said that

$$|\hat{\mathrm{e}} - (\hat{\mathrm{s}} + d_n\hat{\mathrm{g}})| < (d_t + d_n) \tan\left(\sin^{-1}\left(\frac{2\sqrt{3}\sin(\frac{\alpha}{2})}{3}\right)\right), \tag{4.5}$$

or,

$$|\hat{\mathrm{e}} - (\hat{\mathrm{s}} + d_n\hat{\mathrm{g}})| < r, \tag{4.6}$$

Figure 4.10: Covering a plane with reception circles.

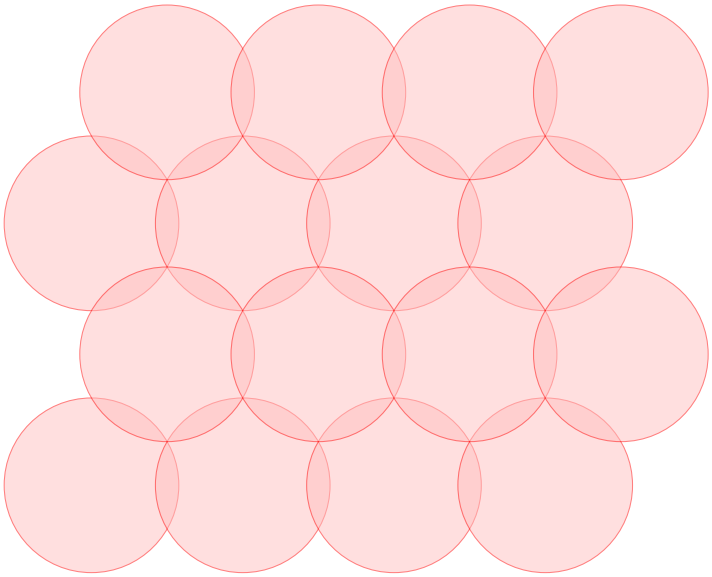


Figure 4.11: The geometry of a ray being received.

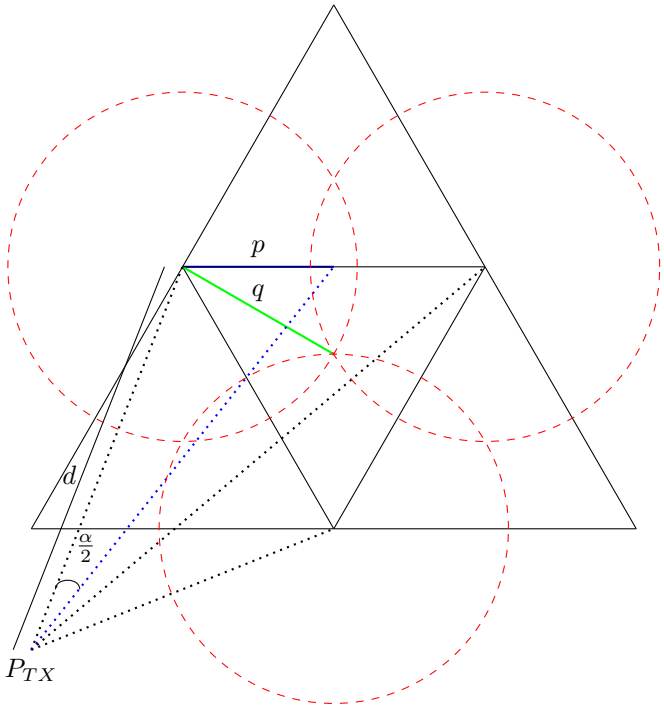


Figure 4.12: Relationship of various values used to determine ray reception.

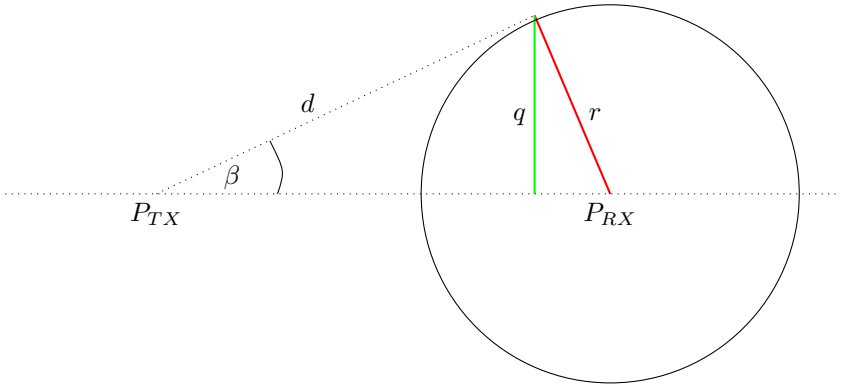
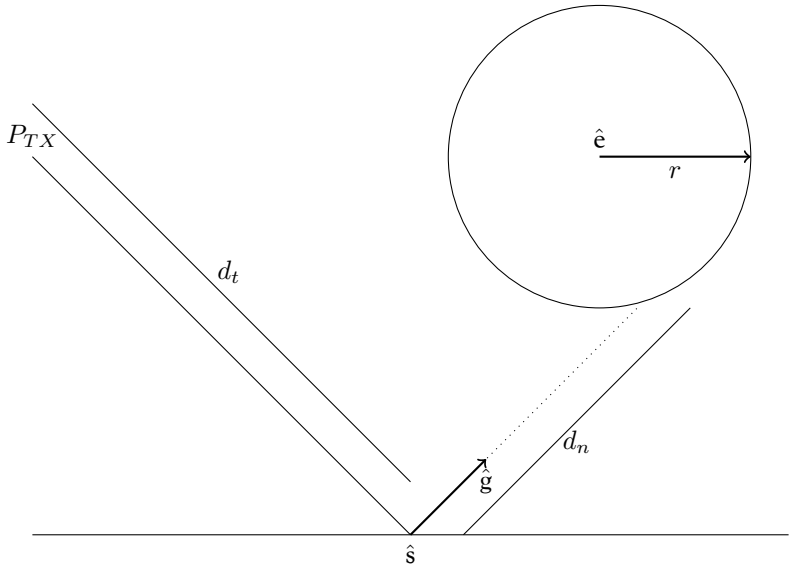


Figure 4.13: A ray intersecting the reception sphere



where $|x|$ means the magnitude of vector x, iff the ray enters the reception sphere at some value of d_n. If the inequality is set to an equality relation, a quadratic in d_n is obtained. The roots of this give the intersection points, if any, with the reception sphere. Note that the sphere is variable in size with d_n, hence the quadratic relation. Rearranging gives

$$d_n^2[\hat{\mathrm{g}}.\hat{\mathrm{g}} - r^2] + d_n[-2\hat{\mathrm{e}}.\hat{\mathrm{g}} + 2\hat{\mathrm{s}}.\hat{\mathrm{g}} - 2d_t r^2] + [\hat{\mathrm{e}}.\hat{\mathrm{e}} - 2\hat{\mathrm{e}}.\hat{\mathrm{s}} + \hat{\mathrm{s}}.\hat{\mathrm{s}} - d_t^2 r^2] = 0. \tag{4.7}$$

In the programme, the discriminant of the quadratic is taken. If this is greater than zero the solutions are computed. These should both be positive for a ray to be received. If one root is positive and one negative, the origin point is already in the sphere, which should never happen, causing an error message to be displayed. If both are negative there is no intersection. The minimum solution is then returned, which is the distance at which the ray segment first intersects the reception sphere. If there is no intersection, because the ray misses the RX an propagates out into space, or because it has hit something else first, then no distance is needed. It should be noted that the

sphere expands between the first intersection and the second, and therefore has a different radius for each. This is because sphere size is variable with d.

4.4 Processing Found Ray Paths

As described in the previous section, ray segments are initialised, traced through the environment and may arrive at the receiver forming a valid ray path. In this process, apart from when deciding the directions of refracted rays, there is no benefit to performing costly electromagnetic calculations 'on the fly'. Most ray paths will not be a valid path between transmitter and receiver; any on the fly electromagnetics for these rejected paths would waste of computational resources.

However, for those ray paths that are found to be valid, various processing needs to be performed. The recursive path search programme stores material types, interaction-locations and what kind of interactions have happened along the route in a stack like structure. This means when a ray path is found to be valid, the processing function can simply process the path using the values in the stack. The processing of a path is the topic of this section.

4.4.1 Ray Processing Using Jones Calculus

Electromagnetic waves are transverse waves which means that the direction of oscillation of the field is perpendicular to the direction of propagation. The oscillation may be in any direction perpendicular to the direction of propagation.

For an electromagnetic wave impinging upon an object, many models exist to predict the behaviour of the interaction, whether this interaction is a reflection, diffraction, refraction or diffuse scatter. Often these models are formulated in terms of two polarisation components. These components are perpendicular to one another. There is usually some concept of a locally horizontal component and locally vertical component. Sometimes the terms 'hard and soft' or 'parallel and perpendicular' are used as alternatives to horizontal and vertical. Note that unless in special circumstances, these local directions do not correspond with any global notion of vertical and horizontal, indeed they may be completely opposite or in any misalignment. For the purposes of this report, a vertical and horizontal concept related to a global notion of "up", the positive Z direction, is used occasionally, especially when first resolving the polarisation of a wave emitted from or arriving to a rotated antenna. In this case, the vertical polarisation direction is perpendicular to the direction of propagation of the ray, and exists in the plane containing both the ray direction vector, and the global Z direction vector. The horizontal direction is perpendicular to the vertical direction and the direction of propagation of the ray.

When a ray leaves the TX, the polarised signal emitted from the antenna is resolved into vertical and horizontal components in relation to a global coordinate system as described above. The vertical direction, $\mathbf{v}_p$, is given by (4.8). The horizontal direction is then calculated to form a left

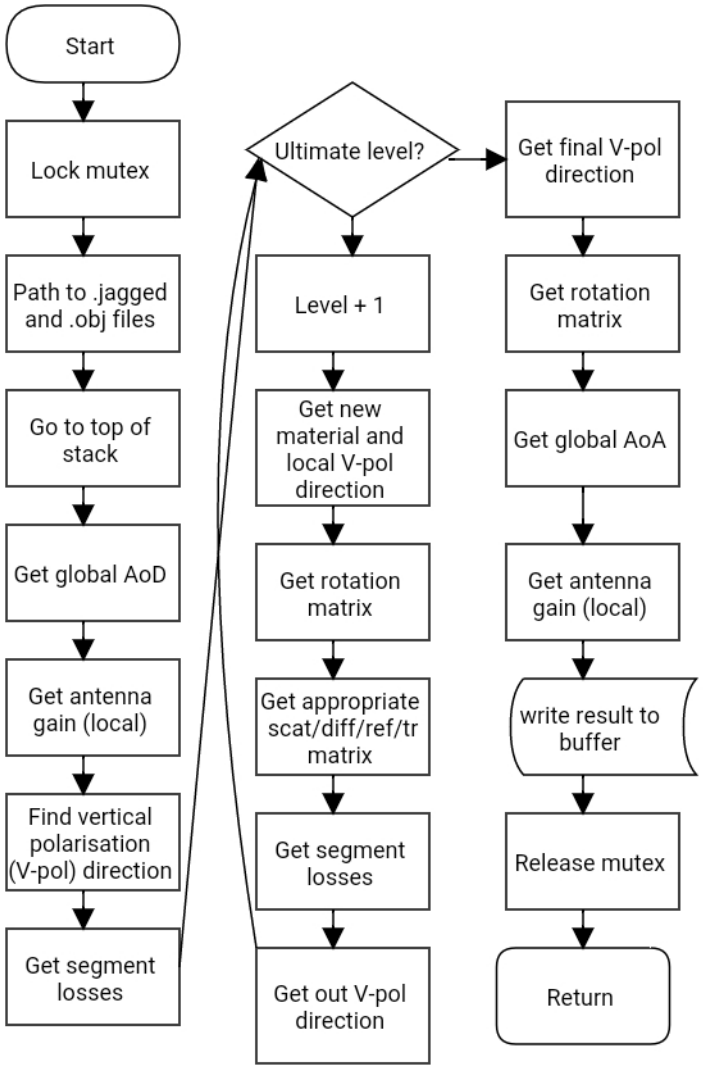


Figure 4.14: Processing a valid ray path.

handed set with this vertical direction and the direction of propagation of the ray.

$$\hat{\mathbf{v}}_p = \begin{bmatrix} -\sin(\theta)\cos(\phi) \\ -\sin(\theta)\sin(\phi) \\ \cos(\theta) \end{bmatrix}. \tag{4.8}$$

Upon the interaction of the propagating ray with an object, the polarisation components are each resolved into local vertical and horizontal components for the incoming ray. This is achieved using a rotation matrix R_{rot}. The local polarisation directions are calculated using the direction of propagation of the ray, and information on the orientation of the object itself. Using the signed angle between in incoming vertical polarisation vector and local vertical polarisation vector around the direction of propagation, α_v, the rotation matrix may be calculated using (4.9).

$$R_{rot} = \begin{bmatrix} \cos(\alpha_v) & \sin(\alpha_v) \\ -\sin(\alpha_v) & \cos(\alpha_v) \end{bmatrix}. \tag{4.9}$$

The incoming local vertical polarisation direction is not, in general, the same as the outgoing local vertical polarisation direction following an interaction. How outgoing polarisation directions are calculated is described in detail in the next subsection of this chapter. Having found the outgoing polarisation direction, this may be stored, and used to calculate the rotation matrix before interaction with the next primitive or receiving antenna. In the case of the receiving antenna, (4.8) may be used, but special attention must be payed to calculating ϕ and θ since the direction of propagation is now towards the antenna rather than away from it. In this way, polarisation may be tracked from the TX to the RX.

To account for feeding an antenna with a particular polarisation, for example, from a waveguide, a two element vector is set. For example, if for a particular simulation vertical excitation of the transmitting antenna was to be simulated, a value of the two element vector s_{tx} of $[1, 0]$ could be used. At the receiver, both permutations of s_{rx} are always used, and then results for horizontal and vertical polarisation are stored in separate .datum files. Particular attention must be given to when this multiplier is used, for example if an antenna is rotated, it is likely that a wave guide would have been rotated with it, rather than feeding globally vertical polarisation in.

Jones calculus, as shown in (4.10), is a powerful tool for computing ray optical power and phase at a receiver. The field delivered to the port of the receiving antenna, E_{RX} is related to the transmitted field, E_{TX}, by

$$E_{RX} = E_{TX} \times s_{TX} \times G_{TX} \times R^0_{rot} \times I^0 \times \cdots R^{n-m}_{rot} \times I^{n-m} \times \cdots \times G_{RX} \times PL. \tag{4.10}$$

In (4.10), $G_{TX/RX}$ is the matrix specifying the transmitter/receiver antenna gain, R^n_{rot} is the rotation matrix for the n^{th} ray segment along a path, I^n is the n^{th} interaction matrix, which could be a reflection matrix, refraction matrix, scattering matrix or diffraction matrix. PL is a loss term. It

includes losses by free space propagation, but also any attenuation due to propagation through a material that has some conductivity.

4.5 Optimisation

Several optimisation techniques were used.

4.5.1 Parallelism

Ray tools typically require the same small set of operations to be performed many times. In the case of the built tool, initial ray directions were decided, and thereafter each one needed go through the same process in which intersection with the environment is tested for and, if the environment is hit, a new ray or rays is generated and then this ray or rays are checked in the same way etc. Because this same operation must be applied for every initial ray in exactly the same way, the method lends itself to parallelism.

The developed tool was designed to run on a Dell Desktop machine with an 8 core CPU. For up to eight initial rays a new thread was created. The recursive path finding algorithm was a depth first search, so sufficient memory was available to run eight threads. There was a resource sharing hazard when writing valid ray paths to a shared result buffer, and then to output files. Access was controlled to this resource using mutual exclusion tokens.

Because not all eight cores were assigned to the application by the operating system, and because the mutual exclusion sometimes causes wait states in concurrent threads, the observed increase in execution speed was not eight times as might be expected, but still significant.

4.5.2 Library Use

The tool made use of two software libraries. The use of the libraries meant that both ray scene intersection tests and matrix operations were handled in an efficient way.

Eigen

Eigen is a matrix library. There are many cases in which matrices are used in the programme, including the processing of ray paths, rotating antenna patterns and the storing of large data structures such as routes. The use of the library means that matrix operations are highly optimised and that these sections of the code are handed efficiently.

Embree

Embree is a software library developed by Intel. It is designed for use in computer graphics applications, where ray tracing is also used as a rendering technique. The library performs extremely

fast ray-scene intersection test, which are repeated continually in the ray launching application. Bounding volume hierarchy acceleration structures are implemented in the library. Once a scene is parsed from an .obj file, it is committed to the embree acceleration structure and from then on ray-scene intersection tests can be very fast.

4.5.3 Tree traversal

The control flow of the programme is recursive in s depth first structure. A ray is launched, it hits something and one or many new rays are generated as appropriate, the recursive function is called on the first of these and so on until the maximum depth is reached, which is limited by the maximum number of interactions. Then the second generated ray in the layer above has the function applied. This depth first search means that the required memory is limited. If a whole tree structure had to be remembered at once, it is likely that most modern machines would fail due to the unavailability of RAM.

4.5.4 Buffering Outputs

When a new ray path is found, it has to be recorded in the output files. Writing to the file-system is a costly operation, and because typically many ray paths will have to be found per step in the route, and because there will be many steps in the route this can aggregate to the dominant operation in the whole application. The solution to this problem is to store results to internal buffers until some maximum buffer size is reached, at which point the buffer is flushed to the output files, with a minimum of overhead, and the programme continues. The programme flushes all the buffers on completion.

4.6 Angular Resolution Analysis

For most scientific models, there is a compromise to be made between complexity of the model and the ability of the model to accurately imitate, in simple terms or concepts, the physical process it represents. Ray launching is used to give insight into the radio channel behaviour in a particular environment at some "cost" less than conducting real-world channel measurements. As it is a software model, various parameters may be set to adjust the accuracy and precision of predictions, usually at the cost of computation time. One example would be turning on modelling of a certain propagation mechanism, diffraction say, which increases computation time significantly, but makes coverage loss predictions in a city more like those observed in real life, especially in NLoS situations. A parameter of interest is the number of rays used, which translates into spatial or angular resolution.

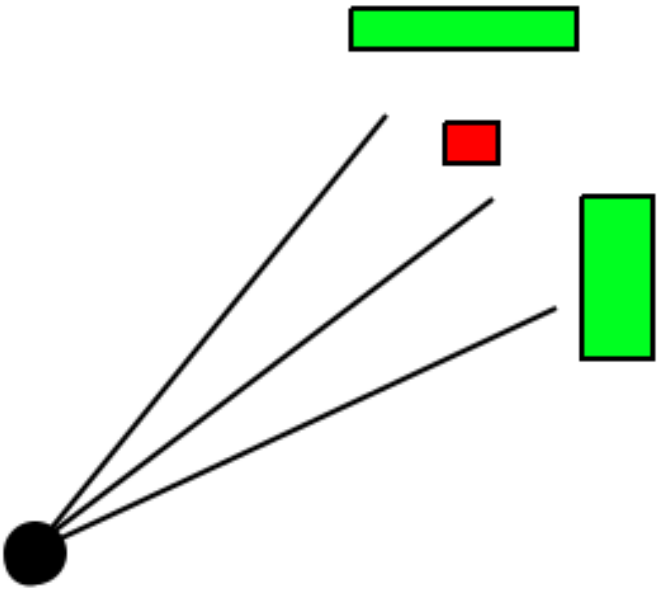


Figure 4.15: missing a small object.

4.6.1 Angular Resolution

The angular resolution impacts prediction accuracy in several ways.

Missing Small Objects

Small items in the environment, and paths through small holes in obstacles may be missed by rays as they diverge from the source as shown in Fig 4.15. Conservatively, the ray launcher must use sufficient resolution such that - after as many allowed bounces across the largest dimension of the environment - it may still hit with at least one ray the smallest surface in the environment. In practice, this is almost impossible and rarely necessary. Most significant reflectors in an environment are large, and for those that are small, unless such a surface is very close to the TX or the RX, their contribution to the signal received at the receiver will be relatively small. Furthermore a path that bounces from corner to corner of a environment that then hits the smallest object in the scene before it arrives at a receiver is unlikely in all but deliberately contrived examples. Generally, the *effective* environment is smaller than a whole environment, as modelled, because of limiting obstructions. For example, in an urban street canyon, most electromagnetic waves that don't propagate up into space will be confined to the street by reflections from buildings, the limiting of the effective size limits the angular resolution needed. Furthermore, the problem of missed paths will be more extreme for long paths, as rays diverge with distance. Because propagation losses are proportional to the inverse of distance squared, this means missed paths are likely to be weak contributors to the received signal. However, in some cases, if an unacceptably low resolution is used, important paths may be missed.

The problem associated with missing small objects can be alleviated by using environment driven ray tracing techniques in which the environment is divided into tiles, but tiles must be made small enough to fit onto small objects in this case.

Path Length

The angular resolution affects the calculated path length. This is because received paths must fall within a reception sphere to count as received, as no rays will exactly intersect an RX location due to the discrete ray launching process. This means as the path length increases, the radius of the sphere needed to capture at least one ray, and therefore the error in path length, also increases. It may be helpful for the reader to refer to figures earlier in the chapter, for example Fig 4.13. Using a lower angular resolution means that the error will be larger for a given path length. Using environment driven ray tracing can alleviate the small-object missing problem, however the path length problem is not solved using these techniques. Because a tile is of finite dimensions and not infinitesimal, firing a ray to its centre also introduces a path length error of a similar kind.

This has two effects. One is a small error introduced in path loss. The model is relatively resilient to this. If a 1 m error is introduced for a path length of 10 m, even at 300 GHz where propagation losses are more acute, there is only an error in received power of about 1 dB, which is typically well within the expected accuracy for a tool of this type. The other is in the prediction of phase at the receiver. This type of prediction may or may not be important. In a coverage prediction scenario, average received power is typically the only metric of interest, and thus accurate phase information is not required. Furthermore, if the environment is multipath rich, each component may be assumed to have a random uniform phase producing Rayleigh fading statistics, which is a reasonable assumption. However, in certain situations, it is highly important to obtain the correct relative phase of ray paths. An example of this would be in order to accurately model fast fading effects where there are a small number of dominant ray paths. Another example is predicting the phase and power received over an array of antenna elements, where it is important to model precise phase relationships between elements for simulation of beam-steering algorithms.

4.6.2 Ray Launching and Analytic 2-Ray Model Above a Flat Earth

Introduction

This section outlines a study comparing a ray launching simulations with a two ray analytic propagation model (implementation of the two ray model part of the "MIMObit" suite [3]). Comparing the ray launching tool to a classical analytic model allows the artefacts introduced by the launching technique to be highlighted. This section is not intended as a verification of the ray launching tool in complex environments, but as an analysis of errors introduced by angular discretisation. The two ray model gives almost instantaneous results for propagation above a flat earth, and this can be

compared to the ray launching results to obtain the accuracy at different angular resolutions.

The two plane wave decompositions, i.e. the ray launcher derived and the analytic, were used in MIMObit to describe scenarios of simple half-wave dipoles, arbitrarily oriented above the ground. The study was limited to evaluating just the total received power in a Single Input Single Output (SISO) setting.

Geometry Description

A scenario was generated in which a TX transmitted to a grid of RXs above a flat earth in the XY plane. The TX was at position (0, 0, 10) m. The RXs were arranged in a 5 by 5 grid with 3 m between RXs in both directions. The central RX in the grid was at position (-1.5, -1.5, 1.5) m. The grid parallel to the XY plane. Each location had a dipole pattern applied. A Z oriented dipole was rotated by by the Euler angles (30, 60, 0) degrees for the TX, and (0, 90, 0) degrees for all the RXs. Thus, the half-wave dipoles of all the RX radios were parallel to the Y axis and they were all 1.5 m above the ground. The ground was given the properties $\epsilon_r = 2.5$ and $\tan(\delta) = 0.02$.

Results and Discussion

The same scenario was simulated twice, once at 3.5 GHz and once at 100 MHz. Because the wavelength was much longer for the 100 MHz simulations, far fewer rays needed to be launched to achieve the same phase error, i.e. the same path length error as a proportion of a wavelength was smaller. This meant that only twenty thousand rays needed to be fired for the 100 MHz simulation to achieve nearly identical power predictions to the analytic model when a coherent sum was used to combine the LoS and reflected ray at each RX location. To achieve the same one eighth of a wavelength phase accuracy, which was found to be that necessary to achieve the acceptable results, six million rays were required for the 3.5 GHz simulations. This simulation therefore took considerably longer, in the order of hours rather than minutes.

When the lower resolution was used for the 3.5 GHz simulations, phase errors that were much greater than a wavelength were introduced, and hence phase was essentially random uniform for both multipath components. This meant that comparing a coherent sum at each receiver to that derived from the analytic results, arbitrarily large errors were introduced.

Appendix A

Example Files

Here, some example input and output files are given for reference.

Listing A.1: A .jagged file

```
1431.000000 0.000000 3.500000 1100.000000 -15.600000 6.000000
+0.058033 -1.495319 0.075279 -0.995474 0.058000 +0.001103 +3.078821 -0.998030
    0.062730 0.001103 1431.000000 0.000000 3.500000 +0.000000 1433.766113 -36.578533
    5.631200 +36.744930 1100.000000 -15.600000 6.000000 +336.428009
+0.030005 +2.738053 -0.919263 0.392499 0.030000 +0.000415 -2.961640 -0.983852
    -0.178983 0.000415 1431.000000 0.000000 3.500000 +0.000000 1357.726563 31.285660
    5.891266 +79.708977 1100.000000 -15.600000 6.000000 +309.685333
```

A jagged file has variable length lines (shown with text wrapping here) with sets of three coordinates specifying x, y and z coordinate in meters of each bounce on a route between transmitter and receiver (included).

Listing A.2: A .datum file

```
2.000000 -30.000000 -1.000000 0.000000 -10.000000 -15.000000
25.548393 85.220266 noFol -150.590512 -237.709552 -2.621080 0.000000
25.533140 85.169387 noFol -150.586779 -228.259964 1.367518 15.452526
25.572536 85.300799 noFol -150.602385 -243.300125 -0.295602 -29.418259
```

A datum file has a special first line with transmitter and reciver localtion in xyz format. Then each line represents a ray path with the following entries: distance (m), time (nS), foliage losses (dB), free space path loss (dB), total loss inc interaction losses (dB), phase shift (rad), Doppler shift (Hz).

Listing A.3: A route file

```
-5, 1, 1.6,      1.5, 2.0, 1.6,  0, 0, 0,       0, 0, 3.141592653589793238462643,  0,
    0, 0,       0, 0, 0
-5, 1, 1.6,      1.5, 1.99, 1.6,         0, 0, 0,       0, 0,
    3.141592653589793238462643,  0, 0, 0,       0, 0, 0
```

```
-5, 1, 1.6,      1.5, 1.98, 1.6,          0, 0, 0,       0, 0,
    3.141592653589793238462643,  0, 0, 0,       0, 0, 0
-5, 1, 1.6,      1.5, 1.97, 1.6,          0, 0, 0,       0, 0,
    3.141592653589793238462643,  0, 0, 0,       0, 0, 0
-5, 1, 1.6,      1.5, 1.96, 1.6,          0, 0, 0,       0, 0,
    3.141592653589793238462643,  0, 0, 0,       0, 0, 0
```

A route file is xyz transmitter location followed by xyz receiver location on each line.

Listing A.4: A materials file

```
0, 1.0, 0.0, 5, 1, 0.0003, 0.1, 0.1, 0.0, 0, 0 ,0
1, 1.0, 0.0, 5, 1, 0.012, 0.0610, 0.1, 0.0, 0, 0 ,0
2, 1.0, 0.0, 3, 3000, 0.00000041, 0.063, 0.1, 0.0, 0, 0 ,0
3, 1.0, 0.0, 5, 1, 0.0003, 0.1, 0.1, 0.0, 0, 0 ,0
4, 1.0, 0.0, 3, 3000, 0.00000041, 0.063, 0.1, 0.0, 0, 0 ,0
```

A materials file has the following entries: material number, real permittivity, imaginary permittivity, real permeability, imaginary permeability, standard deviation of surface roughness (m), correlation length (m), further entries varied during the course of the PhD project and could include speed of an object of that material or different roughness parameters.

Listing A.5: A macro file

```
TXV:60GHzV
TXH:60GHzH
RXV:60GHzV
RXH:60GHzH
rayNum:50000
recDepth:3
frequency:60000000000
sTx:1,0
objFile:corn
reloadObjFile:y
databaseFile:database
routeFile:route
unit:m
doTrans:0
doScatter:n
doDiffrac:y
scatMod:k
```

Listing A.6: A stats file

```
mean excess delay(ns),  9.00181e-006
rms delay spread(nS), 0.00507192
coherance bandwidth(GHz) , 197.164
Doppler spread (Hz), 2.22507e-308
```

```
rms Doppler spread (Hz), 0
coherance time(s) , 4.49423e+307
AoD elevation rms(rad) , 0.000867249
AoD azimuth rms(rad) , 3.98425e-005
AoA elevation rms(rad) , 0.000867249
AoA azimuth rms(rad) , 4.36157e-005
```

Listing A.7: A materials lookup file

```
None, 0
Brick, 1
Concrete, 2
```

This file associates a material name derived from modelling software to a materials property index as specified in the materials file. This allows a human readable name to be associated with a set of electromagnetic properties.

Bibliography

[1] B. S. Lee, A. R. Nix, and J. P. McGeehan. Indoor space-time propagation modelling using a ray launching technique. In *Antennas and Propagation, 2001. Eleventh International Conference on (IEE Conf. Publ. No. 480)*, volume 1, pages 279–283 vol.1, 2001.

[2] E. Mellios, G. S. Hilton, and A. R. Nix. Ray-tracing urban picocell 3d propagation statistics for lte heterogeneous networks. In *2013 7th European Conference on Antennas and Propagation (EuCAP)*, pages 4015–4019, April 2013.

[3] NEBENS. Mimobit. `http://www.nebens.com`, 2018.

[4] S. Y. Seidel and T. S. Rappaport. A ray tracing technique to predict path loss and delay spread inside buildings. In *Global Telecommunications Conference, 1992. Conference Record., GLOBECOM '92. Communication for Global Users., IEEE*, pages 649–653 vol.2, Dec 1992.

[5] Scott Y Seidel and Theodore S Rappaport. Site-specific propagation prediction for wireless in-building personal communication system design. *IEEE transactions on Vehicular Technology*, 43(4):879–891, 1994.

Chapter 5

Diffuse Scattering and Diffraction

Some results in this chapter were obtained in collaboration with Alberto Loiza Friere [45][44]. The work [45] is reproduced with permission from EurApp. In this chapter, the submitted work [44] is also reproduced including figures and tables to which the following copyright will apply assuming publication:

5.1 Introduction

Diffuse scattering and diffraction are explored in this chapter. The are considered as separate phenomena, and diffuse scattering is explored only in the sense of diffuse scattering due to microroughness on surfaces. This means the work here is distinguished from that more familiar to the radar cross section community in which a diffuse scatter effect is introduced because of diffraction from the edges of small surfaces, and likewise from models for macro roughness on surfaces.

5.2 The Importance of Diffuse Scattering

Millimetre wavelength wireless links have been proposed as one of the enabling technologies for 5G communications networks [38]. Models for propagation mechanisms specific to millimetre wavelength wireless channels are required for system design and network planning. One such important propagation mechanism is scattering from rough surfaces [30][47]. Whether diffuse scattering occurs due to surface roughness is determined by a relationship between the size of surface irregularities and the wavelength of the incident radiation. In fact, electromagnetically, a surface may be rough for say, visible light, but smooth for radio waves which have longer wavelengths. If bumps and troughs on a surface are of a similar height to a wavelength of the carrier wave, then a plane wave incident on such a surface will have various phase offsets at different points on the wavefront. This is illus-

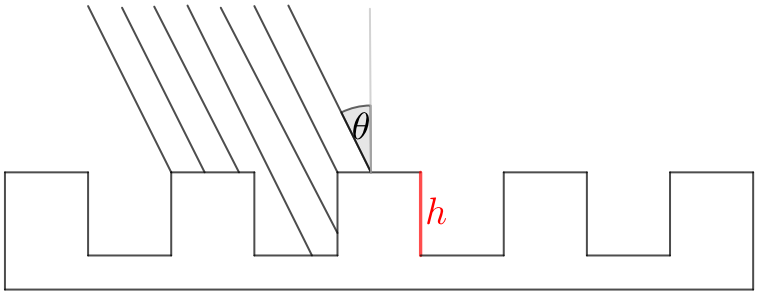


Figure 5.1: A rough surface with an incident plane wave represented as a series of parallel rays.

trated in Fig 5.1. If h is the height of the bumps, and θ is the angle the incident wave makes to the normal of the surface, the Rayleigh scattering criterion, which determines whether as surface may be considered smooth, is given for an incident plane wave of wavelength λ, as

$$\frac{2h\cos(\theta)}{\lambda} < \frac{1}{8}.[8] \tag{5.1}$$

The criterion was somewhat arbitrarily chosen. Rayleigh chose an eighth of a wavelength total extra distance travelled to reach the bottom of a trough and bounce back to the surface before he deemed that the phase difference was significant enough for scattering to begin to occur [8], other fractions have been suggested as more useful in practice [43]. In reality there will be a smooth transition between rough and smooth [8]. The interaction of many reflected components with varying phase relationships means that the incident energy is scattered in nonspecular directions and there is a corresponding reduction in energy scattered in the specular direction [8].

One of the conditions inherent to the use of ray optical techniques is that reflecting surfaces are much larger than the wavelength of the radiation of interest [42]. Clearly with a surface with bumps that are comparable to a wavelength this is no longer true. Furthermore, to model every rough surface in a scene, each with many thousands of facets, would require an enormous level of complexity. It is extremely unlikely that such a model of a large outdoor environment could even be fit into memory on most computing platforms, let alone be processed and used to generate ray tracing data. An alternative used in both the computer graphics and radio propagation communities is to model a simple surface explicitly, and then assign roughness statistics to the simple surface [56][55]. In this way, a rough stone wall might be modelled with a single plane, rather than many thousands or millions of micro facets. This is illustrated in Fig 5.2. The main idea is that the small irregularities cannot be resolved, but the overall effect of these facets on the incident radiation (i.e. diffuse scattering), is still present. A diffuse scattering model determines how much power is scattered from the incidence direction to the scattering direction, dependent on the roughness statistics of the surface it hits.

For computer graphics applications, because of the large number of surfaces that scatter elec-

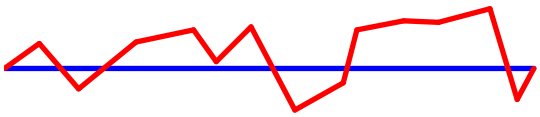


Figure 5.2: A multi-facet composite surface, shown in red, is modelled as a simple surface, shown in blue. Surface statistics derived from the red surface are associated with the blue surface. Several macro surfaces could comprise a street scene with manageable complexity.

tromagnetic radiation at optical frequencies, scattering models have found wide application [55]. Popular models include the Lambertian model [55], Phong model [31], Cook Torrance model [17] and the GGX model [54]. Generally models are investigated to determine if they satisfy reciprocity, conservation of energy and positivity laws [13], although for most applications in computer graphics, visual plausibility is all that is required. It should also be mentioned that at the optical frequencies, many materials display subsurface scattering properties, so the scattering that these models attempt to recreate is not necessarily due to surface roughness alone.

In the radio propagation community, there has been less focus on scattering due to surface roughness. Indeed at sub 6 GHz frequencies most built and natural surfaces are smooth. The phase errors introduced due to extra distance travelled to reach troughs on the surface are much smaller than a wavelength, so reflected components combine constructively, almost as if the surface was perfectly smooth. Scattering has been studied more as a function of the small size of a scattering elements, in terms of radar cross section (e.g. [14]), or due to macro roughness or clutter, e.g. balconies, open windows, street furniture and decorative architectural elements as in [9].

There has been recent interest in adapting scattering models to mmWave applications, as in [30]. There has been a return to more physically based models such as the work in [33]. One of the aims of this chapter is to verify a diffuse scattering model against measurements. Where this has been attempted already in literature [30][12][19][27][16][20], typically complex environments have been modelled, making it difficult to isolate a single interaction for verification purposes.

5.3 Existing Models

5.3.1 Effective Roughness

The effective roughness (ER) model is arguably the most popular model currently used for radio propagation purposes. It's inception was in the idea that macro roughness and clutter was difficult to model over large scale environments like cities [10]. The use of the model meant that features such as balconies, window frames and decorative architectural elements need not be modelled explicitly, and instead their effect on sub 6 GHz incident waves (namely scattering) could be included

in propagation predictions by use of the model. This means the model has a fairly heuristic nature.

The various parameters used to tune the model are generally calibrated against measurements [30][53][22]. These parameters include, choice of a scattering pattern; if a directional pattern is used, how tight or wide the directional beam is and the amount of depolarisation applied [53]. It was mentioned in [10] that some of these parameters may be derived from the standard deviation of surface height of the surface. It isn't entirely clear how this would be done from the references accompanying the statement (i.e. [46][3]). Furthermore it seems only some parameters may be derived in this way, so there is no estimation of depolarisation, main beam width or scattering pattern. In [8], it is shown that correlation length - another important quantifier of roughness and property of the surface - also has a significant effect on scattering behaviour. This means it is unlikely that surface height standard deviation could be used alone to extract model parameters giving realistic results, motivating the need to calibrate against measurements. To calibrate the model against measurements the general procedure is to:

1. Take measurements.
2. Estimate model parameters.
3. Compare to measurements using the fractional index of variance metric.
4. Go to 2) if unsatisfactory performance achieved, else end.

The main disadvantage of this model is the requirement to have measurement data to calibrate the model with. This means the calibrated parameters are also frequency specific, because if a new frequency is used the scattering behaviour will be different, and therefore the old parameters invalid.

In comparison of the ER model to measurement campaigns it is shown that properly calibrated, it performs extremely well in terms of reproducing actual scattering phenomena in built environments. Whilst the model was designed for use at sub 6 GHz frequencies on clutter type roughness, it has been investigated for use in mmWave systems on surface roughness [30][39][12][48][15].

5.3.2 Kirchhoff

The Kirchhoff model for scattering from rough surfaces [8] has been used to model an elementary interaction of an electromagnetic wave with a rough surface at THz frequencies [36][34][21], at 60 GHz to model scattering in an underground mine [49] and to model scattering off asphalt [51].

The Kirchhoff assumption is that the field present at a particular point on a scattering surface can be approximated by the field on the plane tangential to the surface at the point. The validity of this assumption is challenged when the arrangement of the transmitter, surface and receiver and nature of the surface are such that there are features on the surface that stop incident radiation reaching all parts of the surface, or that there are sharp discontinuities on the surface [8].

Whilst it is limited to surfaces with Gaussian surface height distributions, an advantage of the Kirchhoff model over the popular effective roughness model is that it is a deterministic physical model and therefore may be simply be supplied with various physical parameters of the rough surfaces in a scene. Once this is done, frequency may be changed at liberty.

5.3.3 Other Models

Most other models have been produced by the computer graphics community for use in optical visible spectrum applications, like rendering the of 3D models.

Lambertian

The Lambertian model is very simple [29]. Regardless of observation direction, exactly the same amount of power is directed in that direction. A surface with properties predicted by the Lambertian model is not strictly possible to realise in reality. It performs poorly at grazing incidences where otherwise matt materials become reflective (this can be observed on asphalt at glancing angles). There are no parameters to tune, and the model suits very rough surfaces that produce a perfectly diffuse reflection. Many real surfaces that appear to produce a perfectly diffuse refection actually do so by subsurface scattering, not surface roughness. This means the Lambertian model is probably a poor choice for all but the roughest surfaces. Whilst the model is not able to model surfaces that produce a directional scatter well, in situations where there is a high level of uncertainty about surface roughness or amount or location of clutter is high, it has been noted to perform well [11], if only because a it makes a "least wrong" prediction.

Phong

The Phong [32] model was developed on the basis that shiny materials have a very sharp drop off of reflected light as the observer moves away from the specular direction, whereas the light scattered from diffuse materials has a lower but more constant level over direction. The model has several parameters, controlling ratio of specular to diffuse scattering and sharpness of specular drop off.

Torrance Sparrow

The Torrance Sparrow model [52] conjectures that a diffuse reflection can be composed of prefect specular reflection off the many visible micro facets on a rough surface, plus a diffuse term to account for shadowing on the surface or multiple internal bounces etc. The main advantage of the model is that it recreates an experimentally observed phenomenon in which the main lobe of a scattering pattern is not always in the specular direction, especially at grazing incidence.

GGX

The GGX model is a newer model in the same vein as the Torrance Sparrow model [54]. Again a theory of micro facets is used. There is no strong rationale behind the name, it does not appear to stand for anything.

5.3.4 Surface Roughness Properties

Parameters supplied to the models include those quantifying the roughness of a surface. Many of the models have heuristic parameters. The Kirchhoff model, an implementation of which is discussed later in the chapter, uses some physical properties of a surface.

Surface Statistics

The roughness of a Gaussian surface can be described by two statistics. The first is the standard deviation of the surface height, denoted as σ. The second, correlation length, L, is the distance at which the height of points on the surface become uncorrelated by a factor e^{-1}. Intuitively it describes the sort of distance over which surface height changes significantly. These are shown in Fig 5.3. Correlation length is calculated using (5.2). $L = \tau$, where τ is such that $C(\tau)$ falls to e^{-1}.

$$C(\tau) = \sum_{y=1}^{M} \sum_{x=1}^{N} \left[\frac{h_{x,y} h_{x,y+\tau}}{h_{x,y}^2} \right] \tag{5.2}$$

5.3.5 Methods for Obtaining Roughness Statistics

Obtaining accurate roughness statistics is an important first step in verifying scattering models, and later using them. Methods for doing so are presented here.

Lidar Scanners

Electronic surface scanners are available. Construction grade devices may not have adequate spatial resolution for surfaces with intricate roughness, but more precise options do exist. In [18], samples of a Penzance-red-stone wall's surface height were obtained using a Faroarm [1]. Here that data has been post processed to an array of surface height values. A section of the wall, about 8 cm by 70 cm in size, was characterised using (5.2). The processed scan of the wall is shown in Fig 5.4. Values obtained from this process are used later when simulating scattering from a rough wall.

Manual Measurement

An alternative approach is presented in [49]. If the correlation length of the surface is long enough, one may simply use a reference flat plane just off the rough surface of interest, and measure from that to various points on the surface to obtain height samples.

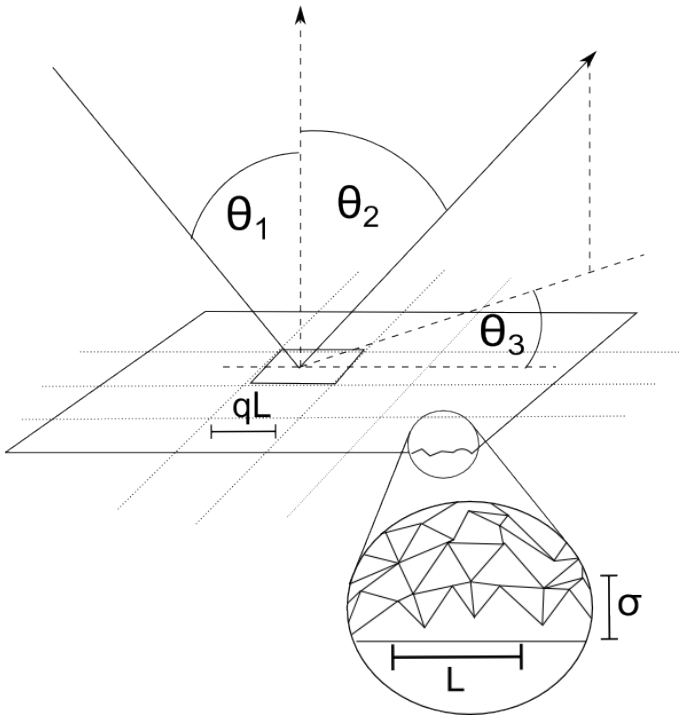


Figure 5.3: Diagram of a reflection, showing how the scattering geometry is defined by three angles. The surface roughness statistics are shown.

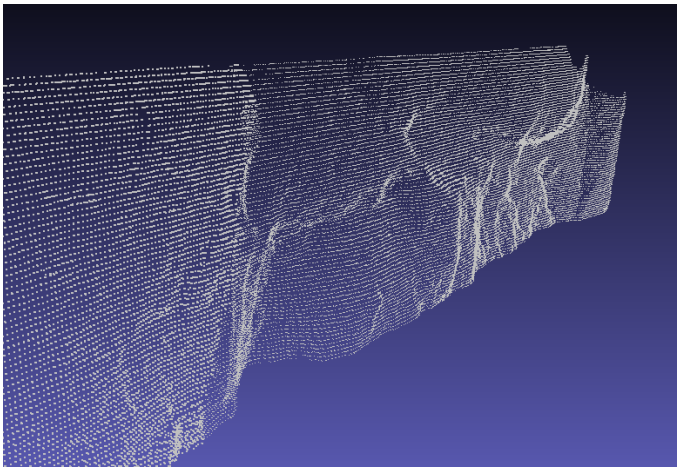


Figure 5.4: Section of wall scan after post processing. Height sample points are separated by 1 mm in the x and y directions.

5.4 Implementation of a Kirchhoff Model

Here, a version of the Kirchhoff model is implemented. The model was selected because it is frequency agnostic. Only one set of parameters need to be defined for a surface then any frequency may be simulated. There are several design decisions that must be made in an implementation, these are discussed.

As stated in the introduction to the model earlier in the chapter, for the Kirchhoff approximation to be valid, the surface may not have sharp discontinuities and there may not be multiple scattering of energy impinging on the surface. To meet these conditions, (5.3) must be satisfied. In (5.3), λ denotes the wavelength of the incident radiation and l_{xy} denotes the length of a side of the square tile to which the model is applied. To meet the leftmost inequality in (5.3), l_{xy} was made to be equal to qL, where q is an arbitrary multiplier greater than 5. Fig 5.3 shows these quantities.

$$l_{xy} \gg L > \lambda \tag{5.3}$$

In [35], it was shown that treating the specular component of an interaction with a rough surface separately from the diffuse component avoids physically unreasonable results due to the choice of the size of the scattering tile to which the model was applied. For this reason, the specular component was de-embedded from the original formulation of the Kirchhoff model [8], leaving only the diffuse component. The specular component was accounted for separately by reducing the standard Fresnel reflection matrix by the roughness correction factor,

$$R_{rough} = R_{Smooth} e^{-g}, \tag{5.4}$$

where g is given in (7.9)[1]. Mathematically the Kirchhoff model for the diffuse components is expressed in (5.5). The model operates in three regimes based on the roughness as indicated by g. The variable $\langle \rho \rho_\infty \rangle$ is the ratio of the average power scattered in a particular direction from a rough tile, to the power of a wave reflected in the specular direction from a perfectly smooth tile of the same size for the same distance and incidence angle.

$$\langle \rho \rho_\infty \rangle = \begin{cases} e^{-g}\left(\frac{\pi L^2 F^2}{A} \cdot e^{\frac{-v_{xy}^2 L^2}{4}}\right), & g \ll 1 \\ e^{-g}\left(\frac{\pi L^2 F^2}{A} \cdot \sum_{m=1}^{\infty}\left[\frac{g^m}{m! m} \cdot e^{\frac{-v_{xy}^2 L^2}{4m}}\right]\right), & g \approx 1 \\ \left(\frac{\pi L^2 F^2}{A v_z^2 \sigma^2} \cdot e^{\frac{-v_{xy}^2 L^2}{4 v_z^2 \sigma^2}}\right), & g \gg 1, \end{cases} \tag{5.5}$$

where

$$F = \frac{1 + \cos(\theta_1)\cos(\theta_2) - \sin(\theta_1)\sin(\theta_2)\cos(\theta_3)}{\cos(\theta_1)(\cos(\theta_1) + \cos(\theta_2))}, \tag{5.6}$$

$$v_z = 2k\cos(\theta_1), \tag{5.7}$$

$$v_{xy} = \sqrt{v_x^2 + v_y^2}, \tag{5.8}$$

with

$$v_x = k \cdot (\sin(\theta_1) - \sin(\theta_2)\cos(\theta_3)), \tag{5.9}$$

[1] In chapter 7 the expression for g is usefully compared to a special case, which is more useful alternative to (5.1) of this chapter.

$$v_y = k \cdot (-\sin(\theta_2)\sin(\theta_3)), \tag{5.10}$$

and A is the area of the scattering tile. The horizontal-horizontal and vertical-vertical Kirchhoff power reflection coefficients, R_{TE} and R_{TM} respectively are then calculated as

$$R_{TE} = \langle \rho\rho_\infty \rangle \cdot r_{TE} r_{TE}*, \tag{5.11}$$

$$R_{TM} = \langle \rho\rho_\infty \rangle \cdot r_{TM} r_{TM}*, \tag{5.12}$$

where r_{TE} and r_{TM} are standard Fresnel field reflection coefficients and $*$ denotes the complex conjugate. Cross polar coefficients are calculated using the perturbation method [37] in order to give a diffuse scattering matrix,

5.4.1 Statistical Variation Around Mean

Statistical variation around the mean scattered power can be generated. For directions that are nonspecular, or for a very rough surface, the distribution of the envelope of the scattered ray is Rayleigh [8]. Since scattered rays were in the non-specular direction, all scattered rays were treated as having a Rayleigh distributed envelope. From [8],

$$P\left(\frac{|\rho|^2}{\langle \rho\rho \rangle} > x\right) = e^{-x}, \tag{5.13}$$

inverting gives

$$|\rho|^2 = \langle \rho\rho \rangle \cdot -\ln(1-p), \tag{5.14}$$

where p is drawn from a uniform distribution with range $[0, 1)$. The phase of the scattered waves is treated as uniform in the range $[0, 2\pi)$.

5.4.2 Sensitivity Analysis

The output of the Kirchhoff model is dependent on various parameters: correlation length; complex permittivity; standard deviation of surface roughness and tile size. The value of q is an arbitrary choice, within the constraints previously discussed. How σ and L may be obtained has

Table 5.1: Sensitivity analysis results

Variable	Range Tested	Δ Peak Power (dBm)
L	30-70mm	3.98
σ	2-8mm	-7.11
ϵ'	1-10Fm^{-1}	1.18
ϵ''	1-10Fm^{-1}	1.59
q	4-20	3.95

been demonstrated in previous sections. It is worth noting however, that complex permittivity is a difficult parameter to measure at millimetre wavelengths. It cannot be inferred from measurements of energy scattered from a surface, since these will include the effects of roughness, not just the permittivity. It is also difficult to measure using standard lab equipment; the use of specialist equipment was described in Chapter 2. Permittivity may vary significantly between sub 6 GHz and tens of GHz [25][50], precluding the use of data obtained at lower frequencies. For example marble has a relative permittivity of 6 sub 6 GHz and a relative permittivity of 10 at 60 GHz. Other materials are more consistent, for example plasterboard appears to remain almost constant over this range [25].

A sensitivity analysis was performed on the model to quantify how changes in input parameters affect the peak output power obtained from a simulation scenario. The simulation scenario is described in the following sections, but briefly, diffuse scattering of incident radiation at 60 GHz was measured in an arc around a rough stone wall. The model was run 120 times with a vector of model parameters chosen at random from a range of plausible values. The results were plotted on scatter diagrams with the dependent variable, peak power, on the Y axis. The independent variable of interest (e.g. the real part of the permittivity) was used on the X axis. Plotting a line of best fit gave an estimate of sensitivity to different parameters. Table 5.1 shows how the peak power varied over the range of plausible input values. One can observe that, in terms of peak power output, the model is most sensitive to the roughness parameters. Complex permittivity is the most uncertain value. From this study, it can be seen that if there is an error of 5 Fm^{-1} in a permittivity value, this will cause an error of less than 1 dB in peak scatted power, suggesting that this uncertainty is not critical.

5.4.3 Comparison to Arc Measurements

To verify the implementation of the model, results predicted from simulations including the model were compared to a measurement campaign conducted at 60 GHz.

Measurement Campaign

The measurement campaign was performed in the main entrance of the Merchant Venturers building (MVB) of the University of Bristol, UK. A transmitter was attached to a pole, 1.7 m from the ground. A receiver was attached to a pole on a trolley and was also 1.7 m from the ground. The trolley was dragged along a 2 m arc around a point on the wall. Both TX and RX antennas were highly directive horn antennas, both were aimed at the point on the wall at the centre of the arc.

The experiment was conducted using channel sounding apparatus used in [7]. A wide-bandwidth, 2 GHz, baseband signal was created and generated using a Keysight M9099 Waveform and Keysight M8190A arbitrary waveform generator (ARB). This signal was taken, in I and Q format, from the direct outputs of the two channels on the device. A 60 GHz carrier signal was generated using a

Figure 5.5: Measurements were conducted in a two metre arc around a section of red-stone wall.

Keysight N5183B MXG Analog Signal Generator. The carrier signal had an actual frequency of 15 GHz, which was fed to a Local Oscillator (LO) port on a Sivers-IMA transceiver, which was used as an up-converter. The transceiver multiplied that frequency by four, to 60 GHz. The transceiver also modulated the 60 GHz signal with the 2 GHz baseband signal. The power signal just before the antenna was 14 dBm; the antenna was fed with a vertically polarised wave-guide. The antenna used for transmission was a high directive circular horn antenna with a HPBW of 12 degrees and a 25 dBi gain. At the receiver, the same antenna was used. The received signal passed through an orthomode transducer to split the received signal into co-polarisation and cross-polarisation components. The orthomode produced an isolation between co and cross polarisation of at least 20 dB, and was connected to two Sivers IMA transceivers, which were used to down-convert both co and cross polar 60 GHz signals into I and Q signals. The orthomode provided a loss of 7 dB gain to the received signal. The down-conversion process required a 15 GHz signal in the LO port of the Sivers IMA transceiver, which was generated in the same way as in the transmitter. The transceiver for co-polar signals had a gain of 4 dB to the received signal. This meant, the actual power received from the antenna was 36 dB below the value displayed on the scope. This includes the 25 dBi antenna gain, the 4 dB transceiver gain (co-polar) and the 7 dB orthomode transducer and waveguide gain. Modelled values were adjusted to include these gains and the 25 dBi gain of the transmit antenna.

The two captured signals were analysed and pre-processed using the Keysight MSOS 804A high performance digital oscilloscope. The Keysight 89600 VSA and Waveform Creator channel sounding function operates by repeatedly transmitting a single carrier signal bearing a modulated waveform.

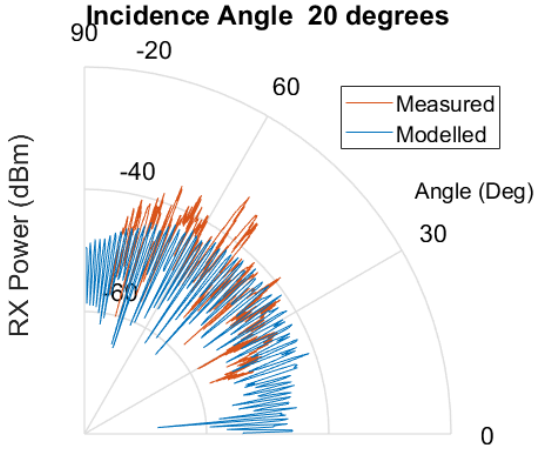
Modelling

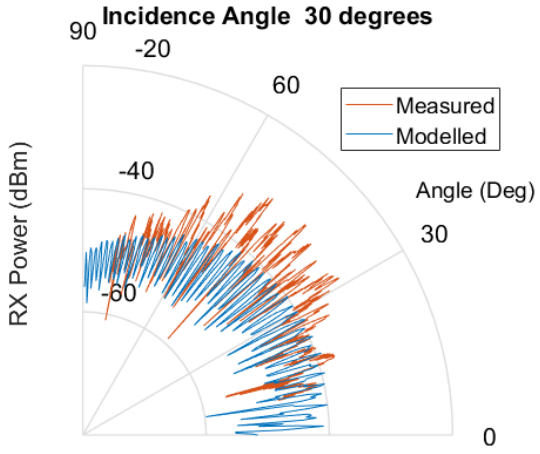
To compare the measurements with a simulation using the implemented Kirchhoff model, the measurement campaign was recreated digitally. The following parameters were used for the wall surface: $\sigma = 3.50mm$, $L = 35.8mm$, $\mu = 1$, $\epsilon' = 9$, $\epsilon'' = 3$, $q = 15$. Roughness parameters were derived from high precision Lidar data. The digital wall surface was divided into tiles to simulate the diffuse contribution from different parts of the wall. If the whole wall was represented by one tile, the multipath characteristic of diffusely scattered energy would be lost. Furthermore the angles that define the scattering geometry of the centre of the tile would be unrepresentative of some regions of the tile generating poor results. The inequality in (5.3) limits the number of tiles used on a finite surface by placing a minimum size constraint upon them, as the tile size must be larger than the correlation length. To match measurements with a certain time resolution, tile size has a maximum size constraint determined by the ability of the receiving apparatus to discriminate between multipath components. If the difference in path length between two paths that are scattered from adjacent tiles causes a difference in propagation time that is greater than the time resolution of the receiving apparatus, then some multipath resolution is lost.

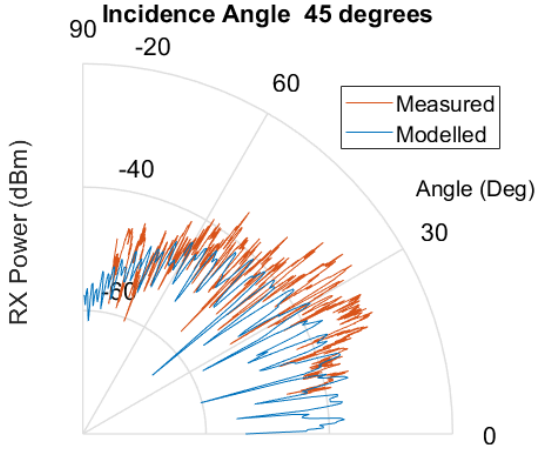
To account for depolarisation, a hybrid model was implemented as in [36], making use of Jones calculus and Kirchhoff/perturbation theory. Co-polar and cross-polar reflection coefficients could be derived for incumbent waves that are perpendicular or parallel to the plane of incidence. This allowed reflected power to be calculated at any position relative to the scattering plane given an arbitrarily polarised incumbent wave. Measured antenna patterns were used to spatially filter ray paths from each tile to accurately replicate a measurement scenario. Since tiles far from the specular direction have low power contributions [34], only 20 horizontal and 10 vertical tiles were used around the point of specular reflection.

Results

Table 5.2 shows the difference in standard deviation and mean between measured and modelled values specified in dB as well as RMSE. Unfortunately none of these are ideal metrics for actually comparing measured and modelled values. Misaligned fades mean that RMSE is likely to be high despite the same fading behaviour (statistically) being produced. The mean says nothing about the scattering pattern, and standard deviation could be a product of fading or scattering pattern shape. The figures depicting measured and modelled values around the arc will are more instructive. It can be observed that the model predicts scattering behaviour quite well for some of the higher incidence angles, but under-predicts power in the specular direction and being most in error for small incidence angles.


Incidence Angle 20 degrees
90
-20
60
-40
-60
RX Power (dBm)
Measured
Modelled
Angle (Deg)
30
0


Incidence Angle 30 degrees
90
-20
60
-40
-60
RX Power (dBm)
Measured
Modelled
Angle (Deg)
30
0


Incidence Angle 45 degrees
90
-20
60
-40
-60
RX Power (dBm)
Measured
Modelled
Angle (Deg)
30
0

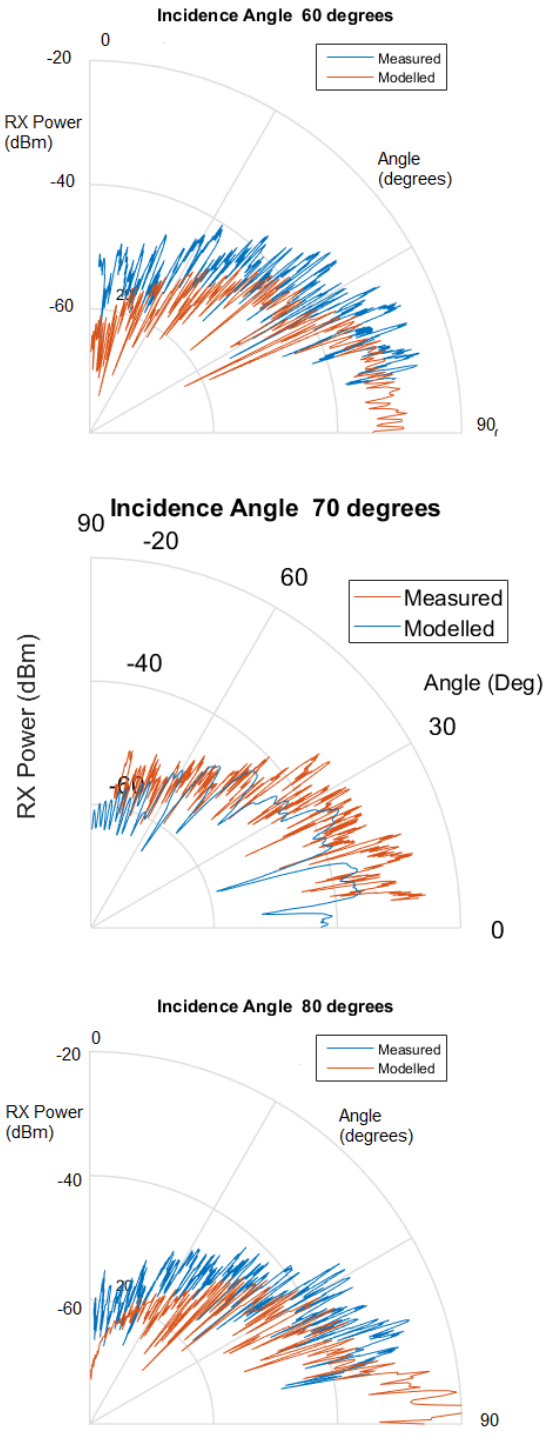


Figure 5.7: Measured and modelled RX power

5.5 A Ray Launching Method for Finding Diffracted Paths

Diffraction is an important propagation mechanism. It is widely reported that for the mmWave bands diffraction is very inefficient, but for sub 6 GHz predictions, it is especially important for NLoS regions.

Diffraction may be included in both ray launchers and ray tracers. When ray tracing, edges are used to create virtual sources on the diffraction cone. When ray launching, for 'environment driven' ray launchers, if a tile is one that borders an edge, it is possible to find edges that create diffracted paths. Here, a method for a ray launcher that uses an angular discretisation at a transmitter point is presented. In [28], a method for finding multiple diffraction edges in a ray tracer is given. Ray launching can scale more favourably with the size of the environment than ray tracing, so finding a robust method for finding diffracted ray paths is of considerable research interest. In [6], the ray tracing and ray launching techniques are combined, but the method by which diffracted paths are found is not described; [41] also uses an unspecified method.

5.5.1 Finding Edges

The process of finding diffraction edges is restricted to those wedges that are oblique and constructed of a non-air material in air. No internal corners are considered, nor are edges existing inside a non-air material. This means that, once a ray has penetrated into a material, no diffracted rays are considered until the ray has propagated back into air.

The algorithm works on a 3D mesh that is composed entirely of triangles, and that has no internal vertices that are attached to only two others, and that has no straight edge with more than one vertex along its length. This means each edge is between two shared vertices of any two tri-

Table 5.2: Measured and Modelled Values Compared

θ_1 (deg)	RMSE (dB)	Δ Std Dev'n (dB)	Δ Mean (dB)
20	6.16	0.74	1.53
30	6.64	1.78	3.81
45	7.09	1.39	3.87
60	8.62	0.44	6.02
70	7.34	0.31	5.13
80	7.45	1.45	5.17
Adv	6.87	0.59	3.21

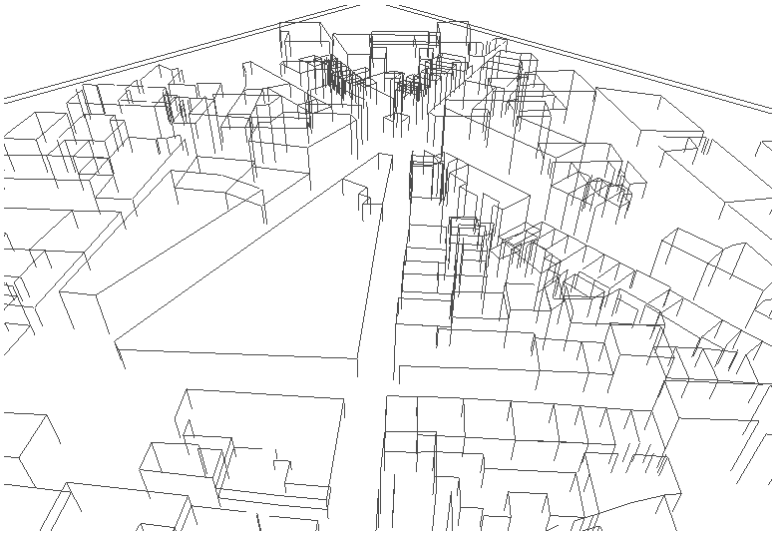


Figure 5.8: A part of London derived with all exterior edges identified. Contains information from open street maps, which is made available at `openstreetmap.org` under the Open Database License (ODbL) `opendatacommons.org/licenses/odbl/`.

angles in the mesh. Whilst this sounds like a complicated set of requirements, it was found that this is the standard output format for the CAD software used to produce environment models throughout this thesis.

All the triangles are compiled into a list. For every triangle in the list, the vertices are checked to be in common with at least two owned by every other triangle below it in the list. If two vertices are shared the normals of the triangles are noted. The wedge angle is equal to $\pi - \zeta$ where ζ is the angle between the normals. A condition on the wedge angle determines if the wedge is an oblique wedge, a small threshold is used to remove nearly flat sections of surface with wedge angles near to π. The normal of the wedge is stored along with wedge angle and the index of the two vertices that the edge is between. Because indices are stored rather than actual vertex locations, this means if the location of the vertices in the scene are updated, due to a dynamic scenario being simulated the edges seamlessly "move" with them.

5.5.2 Ray Edge Intersections

A ray segment can be said to have intersected an edge when the minimum distance between the ray segment and the edge is less than the distance between two diverging rays having travelled to this point. This is measured from the last location of re-spreading (e.g. antenna emission or previous diffraction). The distance to the point along the ray path must also be shorter than the point of any other interaction along the ray path (i.e. the ray must not be blocked).

To find if a ray segment has hit a diffraction edge, and if so which one, it is a simple case of checking every edge, to see if the ray intersects it, and if there is at least one unblocked intersection,

the closest one is used to generate the diffracted rays. For those rays that hit nothing before being checked, the ray is given a very large final distance, to make it into a ray segment, and then this is used to calculate which edges are hit.

To calculate the minimum distance between two ray segments first the two lines are found in vector form [4].

$$R(s) = R_0 + s(R_1 - R_0) \tag{5.15}$$

and

$$E(t) = E_0 + t(E_1 - E_0). \tag{5.16}$$

In the above R is the line containing the ray segment, E is the line along the edge of interest, s and t are scaling variables, and R_1, R_0, E_1 and E_0 are the end and start points of the ray and edge respectively respectively. A vector from any point on R to any point from E may be formed.

$$W(s,t) = (R_s - E_t). \tag{5.17}$$

Finding the minimum distance between two lines is the same as minimising $W(s,t)$ over all s and t. The solution vector, $W(s_c,t_c)$, has the unique property that it is perpendicular to both lines, whereas any other vector connecting a point on E and R will not have this property. This means $R \cdot W_c = 0$ and $E \cdot W_c = 0$. Letting $w_0 = E_0 - R_0$, u be a direction vector for E_t and v a direction vector for R_t, simultaneous linear equations are formed

$$(u \cdot u)s_c = (u \cdot v)t_c = -u \cdot w_0. \tag{5.18}$$

$$(v \cdot u)s_c = (v \cdot v)t_c = -v \cdot w_0. \tag{5.19}$$

Letting $a = u \cdot u\, b = u \cdot v\, c = v \cdot v\, d = u \cdot w_0$,

$$s_c = \frac{(be - cd)}{ac - b^2} \tag{5.20}$$

and

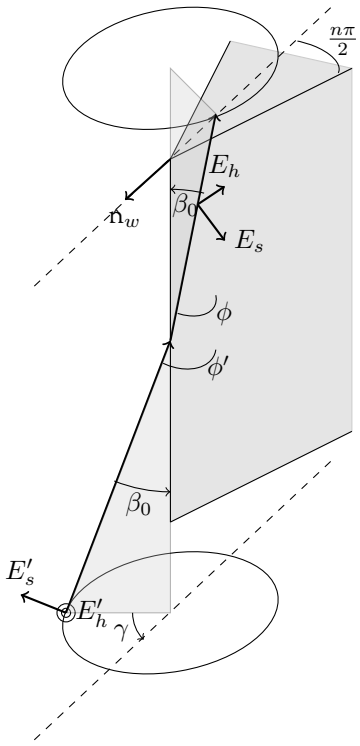
$$t_c = \frac{(ae - bd)}{(ac - b^2)}. \tag{5.21}$$

This gives the closest points on the two lines. However, because the diffraction edges have finite length, and the ray also has an origin and point where it terminates, further calculation must be performed. Minimising $W(s,t)$ is equivalent to minimising the following

$$|w|^2 = (w_0 + su + tv) \cdot (w_0 + su + tv) \tag{5.22}$$

a quadratic in w with a unique minimum. If the minimum lies outside the regions $0 < s_c < 1$ and $0 < t_c < 1$, either the minimum is within the diffraction edge or ray segment but off one of the two ends of the other, or it is off one end of each. This means there are 9 possible cases that must be checked, and s_c and t_c moved to the nearest appropriate boundary of the feasible region as necessary.

Figure 5.9: Diffraction from a wedge



5.5.3 Generating Diffracted Rays

For a ray incident on an edge, β_0 the angle between the ray and the edge. Diffracted rays are generated on a cone with a vertex angle of $2\beta_0$ such that each diffracted ray is also at an angle β_0 t the edge. If the wedge has normals $\hat{\mathbf{n}}_1$ and $\hat{\mathbf{n}}_2$ the wedge angle, w, is given by,

$$w = \pi - \arccos\left(\frac{\hat{\mathbf{n}}_1 \cdot \hat{\mathbf{n}}_2}{|\hat{\mathbf{n}}_1| \times |\hat{\mathbf{n}}_2|}\right) = n\pi \tag{5.23}$$

The "normal" of the wedge $\hat{\mathbf{n}}_w$ is defined as

$$\hat{\mathbf{n}}_w = \frac{\hat{\mathbf{n}}_1 + \hat{\mathbf{n}}_2}{|\hat{\mathbf{n}}_1 + \hat{\mathbf{n}}_2|} \tag{5.24}$$

where the normal describes the direction pointing "out" the point of the wedge. An objective of the design of the code was to have diffracted rays spaced α rads apart. This was to mirror the spreading angle at the transmitter which is also α rads. To arrange rays on a cone such that they are α rads apart, the rays must be rotated around the central axis of the cone by ϵ degrees. For continued analysis a transformation is made such that the wedge edge is aligned with the x axis, and the wedge normal is aligned with the z axis. The n^{th} diffracted ray on the diffraction cone has the direction vector $\mathbf{d}_{rn}$, which can be calculated as

$$\mathbf{d}_{rn} = \begin{bmatrix} \cos(\beta_0) \\ \sin(\beta_0)\sin(n\epsilon) \\ \sin(\beta_0)\cos(n\epsilon) \end{bmatrix}. \tag{5.25}$$

It will be helpful to note that, if an initial ray is constructed such that the ray is in the same plane as both the wedge normal and the diffraction edge, then $n = 0$ and therefore the ray will have direction

$$\mathrm{d}_{r0} = \begin{bmatrix} \cos(\beta_0) \\ 0 \\ \sin(\beta_0) \end{bmatrix}. \tag{5.26}$$

The next ray around the edge would be found where $n = 1$ and therefore

$$\mathrm{d}_{r1} = \begin{bmatrix} \cos(\beta_0) \\ \sin(\beta_0)\sin(\epsilon) \\ \sin(\beta_0)\cos(\epsilon) \end{bmatrix}. \tag{5.27}$$

It is desired that the angle between two rays, as a function of ϵ, be α. Using the dot product to calculate the angle between two ray directions d_{r1} and d_{r0}, which will be the same angle as between all subsequent subsequent pairs of rays,

$$\cos^{-1}(\cos^2(\beta_0) + 0 + \sin^2(\beta_0)\cos(\epsilon)) = \alpha. \tag{5.28}$$

This means that

$$\epsilon = \cos^{-1}\left(\frac{\cos(\alpha) - \cos^2(\beta_0)}{\sin^2(\beta_0)}\right). \tag{5.29}$$

Unfortunately, to distribute multiple rays around the cone, this places some restriction on allowed values of β_0 given an α. The maximum value of α is achieved when $\epsilon = \pi$. This means, to be able to fire more than one diffracted ray from the edge, $2\beta_0 > \alpha$, remembering that $\beta_0 < \pi/2$. If only one ray is fired, which is always possible, results become less meaningful as the chosen direction of the ray, which would be arbitrary, might affect results significantly - should the ray be fired down the front face, back face or out from the edge? In practice, because α will generally be small, only in the case of glancing incidence will this be problematic.

In software, when a diffracted ray hits a wedge, ϵ is calculated, and rays are spaced around the cone starting with one in the plane of the front face. Any rays inside the diffraction wedge, that is if the angle between the ray projected into a plane with the edge as normal, and the wedge normal, is greater that $\pi - w/2$ then the ray is ignored because it is internal. Other rays are free to propagate. A special case is the ray that is nearest to the back face, but internal. This is moved to propagate in the plane to the back face. This ensures subsequent edges on the solid are found. Doing this could create double counts, but filtering of similar ray paths solves this problem effectively.

5.6 Uniform Theory of Diffraction

The uniform theory of diffraction [24] was based upon Joseph Keller's geometrical theory of diffraction [23], but uses transition functions to remove the singularities at shadow boundaries that were present in Keller's formulation.

5.6.1 Diffraction from a Perfectly Conducting Wedge

The soft and hard diffraction coefficients are calculated as

$$\begin{aligned} D_{s,h} = & \frac{-exp[-j(\pi/4)]}{2n\sqrt{2\pi k}\sin(\beta_0)} \\ & \times \Bigg[\cot\left(\frac{\pi+(\phi-\phi')}{2n}\right) F[kLa^+(\phi-\phi')] \\ & + \cot\left(\frac{\pi-(\phi-\phi')}{2n}\right) F[kLa^-(\phi-\phi')] \\ & \mp \Bigg\{ \cot\left(\frac{\pi+(\phi+\phi')}{2n}\right) F[kLa^+(\phi+\phi')] \\ & + \cot\left(\frac{\pi-(\phi+\phi')}{2n}\right) F[kLa^-(\phi+\phi')] \Bigg\} \Bigg], \end{aligned} \tag{5.30}$$

where

$$a^{\pm}(\beta) = 2\cos^2\left(\frac{2n\pi N^{\pm}-(\beta)}{2}\right), \tag{5.31}$$

in which $N^{\pm}$ are the integers which most nearly satisfy the equations

$$2\pi n N^{+} - (\beta) = \pi \tag{5.32}$$

and

$$2\pi n N^{-} - (\beta) = -\pi \tag{5.33}$$

with

$$\beta = \phi \pm \phi'. \tag{5.34}$$

Depending on the type of wave incident on the wedge, L takes different forms.

$$L = \begin{cases} s\sin^2(\beta_0), & \text{if plane incidence} \\ \frac{ss'\sin^2(\beta_0)}{\sin(\beta_0)(s+s')}, & \text{if cylindrical incidence} \\ \frac{ss'}{s+s'}\sin^2(\beta_0), & \text{spherical incidence,} \end{cases} \tag{5.35}$$

where s' is the distance to the diffraction wedge and s is distance from the wedge to the observation point. After the diffraction coefficients have been multiplied by the incident field terms, a spatial attenuation factor, A, is applied which also depends on the type of wave incident on the wedge.

5.6.2 Multiple diffraction from a dielectric wedges

While (5.30) gives the correct dyadic reflection coefficients for a perfectly conducting wedge, in reality most edges encountered will be made of a dielectric material. The following formulation may be used in this case [26]

ϕ'	$\lvert\phi' - \phi\rvert$	θ_0	θ_N
$< \pi$	$< \pi$	$-\phi'$	$-(\phi + \phi')$
$< \pi$	$> \pi$	ϕ'	$n\pi - (\phi + \phi')$
$> \pi$	$< \pi$	ϕ'	$n\pi - (\phi + \phi')$
$> \pi$	$> \pi$	$n\pi - \phi'$	ϕ

Table 5.3: Choice of angle for calculation of reflection coefficients.

$$
\begin{aligned}
D_{s,h} = \frac{-exp[-j(\pi/4)]}{2n\sqrt{2\pi k}\sin(\beta_0)} \\
\times \Bigg[\cot\left(\frac{\pi + (\phi - \phi')}{2n}\right) F[kLa^+(\phi - \phi')] \\
+ \cot\left(\frac{\pi - (\phi - \phi')}{2n}\right) F[kLa^-(\phi - \phi')] \\
+ R_0^{s,h} \cot\left(\frac{\pi + (\phi + \phi')}{2n}\right) F[kLa^+(\phi + \phi')] \\
+ R_N^{s,h} \cot\left(\frac{\pi - (\phi + \phi')}{2n}\right) F[kLa^-(\phi + \phi')] \Bigg],
\end{aligned}
\tag{5.36}
$$

where $R_0^{s,h}$ is the [soft, hard] reflection coefficient at the incidence face, and $R_N^{s,h}$ is the [soft, hard] reflection coefficient for the diffraction face. Unfortunately this new formulation still fails in some cases. Remley proposed new angles to use in the calculation of the reflection coefficients [40]. These are given in Table 5.3. Using these angles removes the restrictions on incidence and diffraction angles present in Luebbers' original formulation.

Choice of Approximation of Transition Function

It is also necessary to compute the transition function, given as

$$F(X) = 2j\sqrt{X}e^{jX}\int_{\sqrt{X}}^{\infty} e^{-j\tau^2}d\tau. \tag{5.37}$$

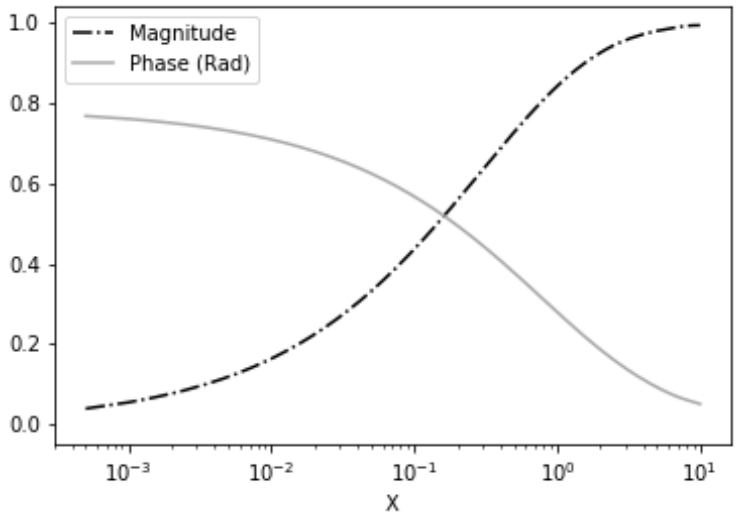
The integral with an infinite limit is not amenable to numerical calculation on a computer. In the original work [24], approximations to the Fresnel integral based transition function are given as

$$F(X) \simeq \left[\sqrt{\pi X} - 2Xexp\left(j\frac{\pi}{4}\right) - \frac{2}{3}X^2 exp\left(-j\frac{\pi}{4}\right)\right] \cdot exp\left[j\left(\frac{\pi}{4} + X\right)\right] \tag{5.38}$$

when X is small,

$$F(X) \sim \left(1 + j\frac{1}{2X} - \frac{3}{4X^2} - j\frac{15}{8X^3} + \frac{75}{16X^4}\right) \tag{5.39}$$

Figure 5.10: Transition function approximation



when $10 > X > 0.001$, and

$$F(X) \sim 1 \tag{5.40}$$

when $X > 10$. Unfortunately comparing these approximations to Fig.5 of that paper, where the transition function is plotted, there is a region between $X = 0.1$ and $X = 3$ where neither (5.38) or (5.39) gives an acceptable approximate result. (16) from [5] is used in place of (5.39) above, meaning for $10 > X > 0.001$,

$$F(X) \simeq \sqrt{\pi X} e^{j(\pi/4+x)} - 2j\sqrt{X} e^{jX} \int_0^{\sqrt{X}} e^{-j\tau} d\tau. \tag{5.41}$$

This reintroduces an integral, which is evaluated numerically in the computer code, but removes the necessity to perform an integral over an infinite range, as would be necessary for (5.37). It was found that (5.38) was better than the suggested alternative, (15) from [5], for $X < 0.001$, so the original formulation, (5.38), was used. A plot of the transition function is given in Fig 5.10.

5.7 Validation

To test the UTD model, predictions were compared to measurements of diffraction around a corner. An experiment was conducted in an anechoic chamber.

5.7.1 Measurements Description

A ninety degree wedge was constructed from MDF and placed in an anechoic chamber. In this way spurious echos from environmental objects were suppressed. An Anokiwave array [2] was used as the transmitting antenna and a square-horn antenna was used as the receiving antenna at 28 GHz. Both antenna were connected to ports on a VNA, which was used to measure the received

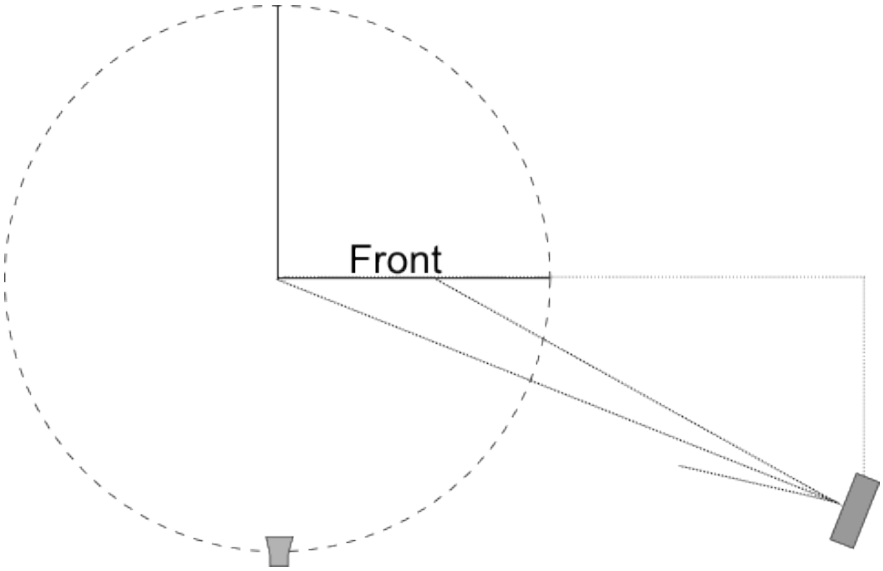


Figure 5.11: Plan view of the experimental set up - to scale.

signal. A birds eye view of the experiment is shown to scale in Fig 5.11. The square horn navigated a circular arc on a robotic arm around the diffraction edge. The radius of the arc was 1.5 m. The arc is indicated in the figure by the dashed circle. The transmitting array was at a twenty degree angle to the front face of the ninety degree diffraction wedge. Its bore-sight was directed at the centre of the diffraction wedges edge. It had a 14 degree 3 dB beamwidth, which is indicated in the figure with dotted lines. It can be observed that a large portion of the front face was strongly illuminated. At the diffraction edge, the boundary of the 3 dB beamwidth illuminated region - in the vertical direction - fell 0.42 m either side of the bore-sight point, which, itself, was half way down the wedge. This means that the top and bottom edges of the wedge - which were 0.61 m either side of the bore sight point - were not illuminated strongly, minimising extra diffracted paths. Along the front face, the array was 3.2 m from the diffraction edge, and was at a perpendicular distance of 1165 mm to the front face.

Two surface finishes were used. The plain MDF was not used, as shown in Fig 5.12. Because the wedge was made from two 9 mm thick sheets, and not a solid block of MDF, it was desired to have a low penetration surface finish so that a single diffraction could be measured, rather than potentially many refracted and reflected paths through the dielectric MDF. To achieve this, aluminium tape was used to cover the wedge. The tape was applied carefully such that the surface could be considered completely flat at the frequency of interest, 26 GHz, this finish is shown in Fig 5.13. As well as the smooth metallic finish, it was desired to have a rough, but still conductive finish. This was achieved by forming aluminium foil over some egg crate foam. Using the form allowed a partially random roughness to be created, which still had broadly the same roughness statistics over the whole surface, see Fig 5.14.

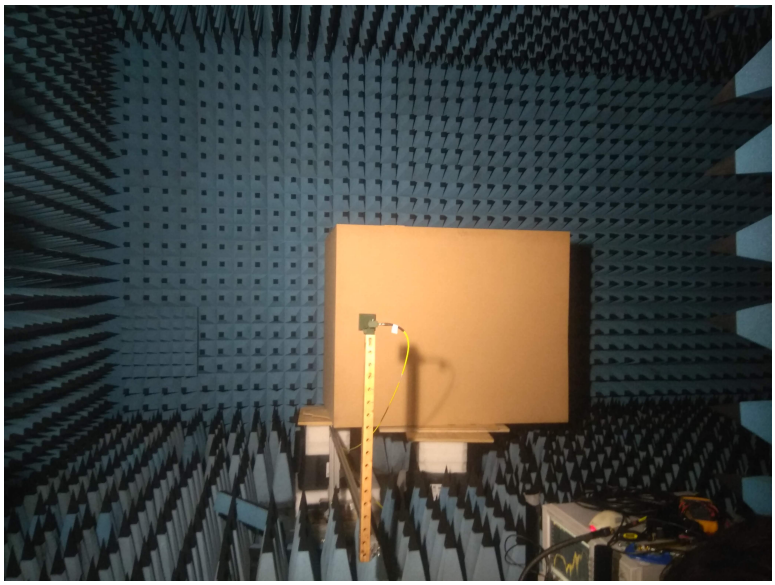


Figure 5.12: The MDF wedge set up in the anechoic chamber.

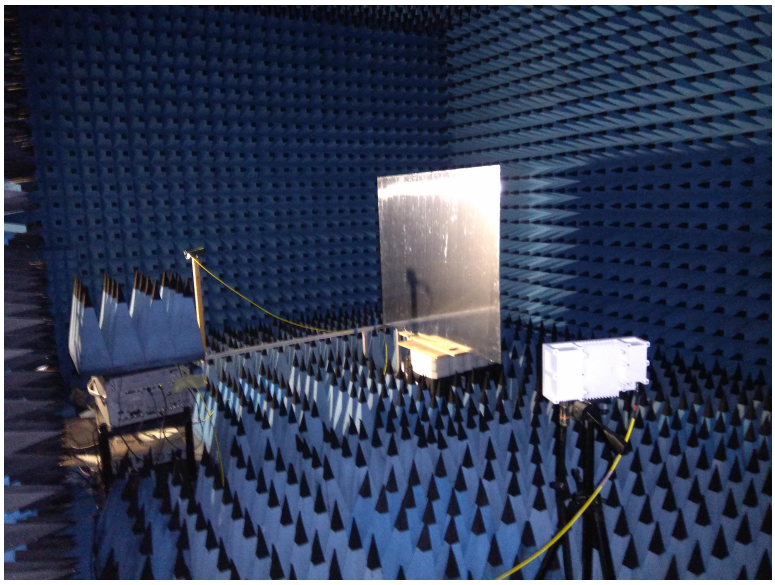


Figure 5.13: The MDF wedge set up in the anechoic chamber covered with aluminium tape to create a conducting wedge.

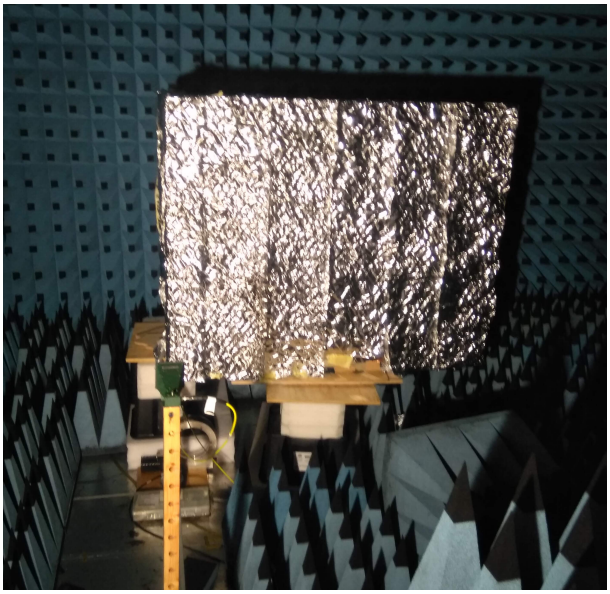


Figure 5.14: The MDF wedge cladded with aluminium foil that was formed over egg crate foam to create a rough conductive wedge.

Robotic Arm

A robotic arm was constructed from aluminium to sweep the receiving antenna through an arc around the diffraction edge. The final upright - i.e. that closest to the receiving antenna - was constructed from wood to minimise unwanted reflections from the arm. The horizontal section was left as bare aluminium. When the horizontal section was covered with absorbing material, it was found that no significant difference was made to the recorded measurement, which was highly repeatable. The absorbing material was removed, as removing weight from the arm made the sweep more stable. The motor driving the robotic arm was a geared stepper motor, and the triggering on the VNA was linked to the signal feeding the motor. Triggering occurred with a 0.1 degree angular sampling frequency, or approximately 2.6 mm spatial sampling frequency - about a quarter of a wavelength at 26 GHz which was the frequency used for the test.

5.7.2 Simulation Description

Simulations were designed to mimic the experimental set up. Fig 5.15 shows a simplified image of the original experiment. Both antennas were measured in the anechoic chamber to extract a radiation pattern that could be used in simulations. The front face was divided into square tiles - as shown in Fig 5.15 - for the scattering model to be applied to. A diffuse scattered ray is shown in red. It hits the centre of a tile on the front face. The tiles had a 12.2 cm edge length. As well as diffuse scattered rays there is a specularly reflected ray, shown in blue, and diffracted ray, shown in hot pink. The robotic arm was omitted from simulations. Before the specularly reflected ray's point of reflection

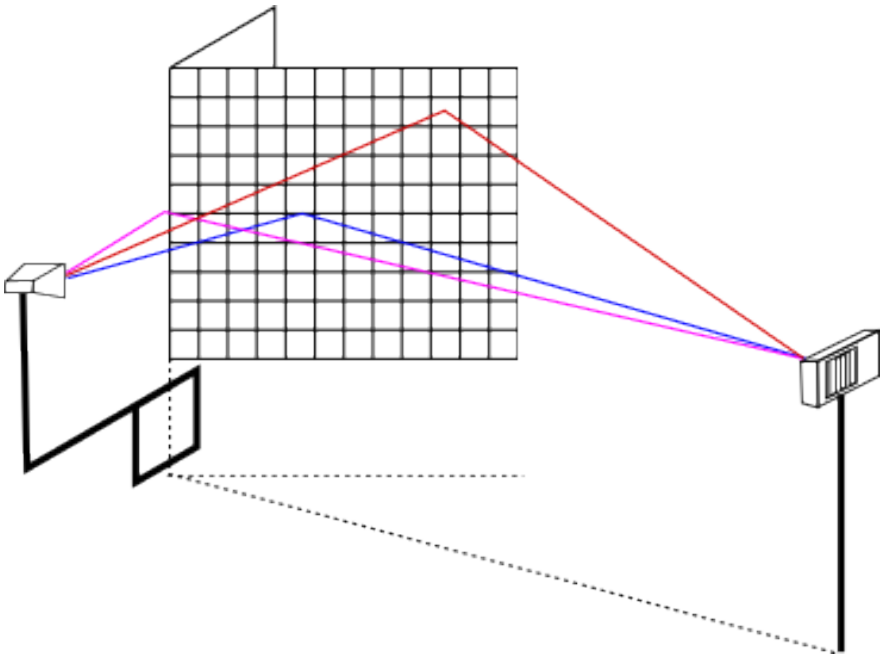


Figure 5.15: Ray paths within 10 dB of the strongest path at 12 m separation to the smooth wall using directional radiation patterns.

Angle from face (deg)	Reflected ray	Scattered rays	LoS ray	Diffracted ray
50 (start) - 110	no	yes	yes	yes
110 - 160	yes	yes	yes	yes
160 - 180	no	yes	yes	yes
180 - 200	no	no	yes	yes
200 - 230 (end)	no	no	no	yes

Table 5.4: Angular range over which different ray types are active.

is on the front face, and after it moved off the end of the wedge, no reflected ray was simulated. Likewise no scattered paths were simulated after the receiving antenna moved behind the front face of the diffraction edge. Finally the LoS ray (not shown) was omitted as soon as the receiving antenna was shadowed by the wedge. The angular ranges over which different ray types were active is summarised in Table 5.4.

For the rough wedge, the roughness properties were obtained using a Faro arm. A processed scan of the manufactured surface is shown in Fig 5.16. The roughness properties are specified in Table 5.5, which were derived from the height data using (5.2) and the usual meaning of standard deviation.

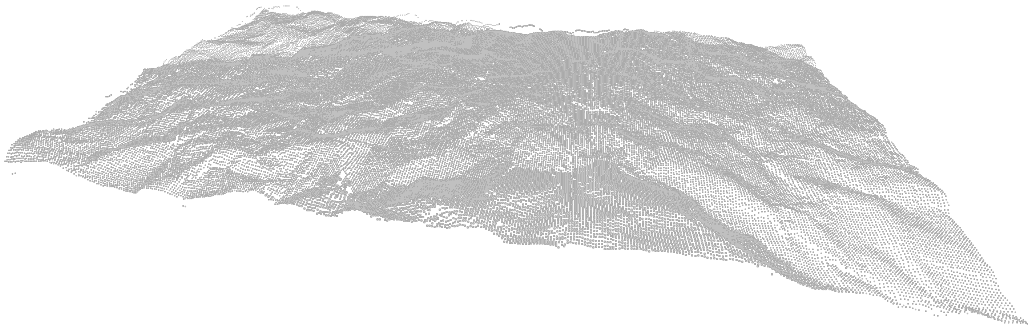


Figure 5.16: Sub-millimetre precision array of surface height values for the manufactured aluminium-foil surface.

σ	L
3.2mm	18 mm

Table 5.5: Roughness properties of aluminium foil formed over egg-crate foam.

5.7.3 Results

A comparison of simulated and measured results are presented in Figs 5.17-5.19. Fig 5.17 shows results for the smooth-surfaced wedge. Figs 5.18 and 5.19 show two different sets of simulated results for the rough wedge. The measured data is the same in both 5.18 and 5.19. Fig 5.18 shows the simulated results using the standard UTD coefficients and Fig 5.19 has them reduced by the Rayleigh roughness correction factor. It can be observed that this makes little difference to predicted results, suggesting the roughness of a wedge does not significantly affect the efficiency of diffraction around it.

It can be observed that - in the umbra region - simulations are not in line with the measured data. Partly, this can be attributed to ray modelling artefacts. The LoS ray transitions from being a valid, very strong propagation path, to being absent from the channel in an instant in simulations. This would be expected to create a discontinuity not seen in measured data, which may be observed in all the simulated results at about 200 degrees. It appears too that the diffracted power is estimated by the model. It is interesting here to compare results to [?]. In that work the authors measured diffraction around a metal wedge at 60 GHz (amongst other related experiments) and although their simulations were confined to a 2D plane, their UTD model did agree with measurements more satisfactorily in the umbra region. It is unclear how they avoided any discontinuity in

Rough	Smooth
1.38	0.928

Table 5.6: FVU comparison for the rough and smooth results.

their simulated received power when the LoS path is suddenly blocked with the metal wedge, but their superior results could also suggest that there may have been some explanatory phenomenon resulting in the discrepancy in this work. Outside the umbra region, it can be seen that predictions are reasonable. Both the LoS and reflected peaks match up well between measurements and simulations. In the simulations for the rough wedge, the depth of fading in the initial section of the sweep where diffuse scattering is the dominant propagation mechanism appears to match measurements well. In the same section for the smooth wedge, power appears to be under predicted. This is the region before the reflected ray comes into existence.

Beyond visual inspection, it is a delicate task to quantify the 'goodness' of a match between measurements and models. This is because there are various metrics of interest that could be extracted from the recorded data. For example, a model that is able to produce statistically similar fading to that found in measurements represents the dynamics of the channel better than one that simply predicts mean received power - with no fades at all - but that has a slightly lower MSQE on average. At other times, minimising error with no special regard for dynamics at all may be a priority. These subtleties will affect how good a model is deemed to be.

A first metric is the fractional variance unexplained (FVU). This has been used to quantify the match between predicted scattered power and measured scattered power in [30]. The metric, because it is scaled by the variance of the measured data, is somewhat scale invariant. Furthermore it does reduce the impact or errors caused by misaligned fades, but it is not perfect in this regard. It is calculated as

$$\text{FVU} = \frac{\text{MSQE}(P_{\text{sim}(dB)}, P_{\text{meas}(dB)})}{\text{Var}(P_{\text{meas}(dB)})} \tag{5.42}$$

and is presented for the rough and smooth wedge in Table 5.6.

A second metric of interest is the mean square error between simulations and measurements where a smoothing average filter has been applied, removing the fast fading present in both data sets. The filter takes the 64 samples either side of a sample of interest and computes a mean. This means at each sample point, recorded data within 16 wavelengths of the point is included in averaging. Filtered data is shown for the smooth wedge in Fig 5.20. MSQE results for the smoothed data for both the smooth and rough wedge is presented in Table 5.7.

Rough	Smooth
5.37	6.77

Table 5.7: MSQE comparison for the rough and smooth results (dB).

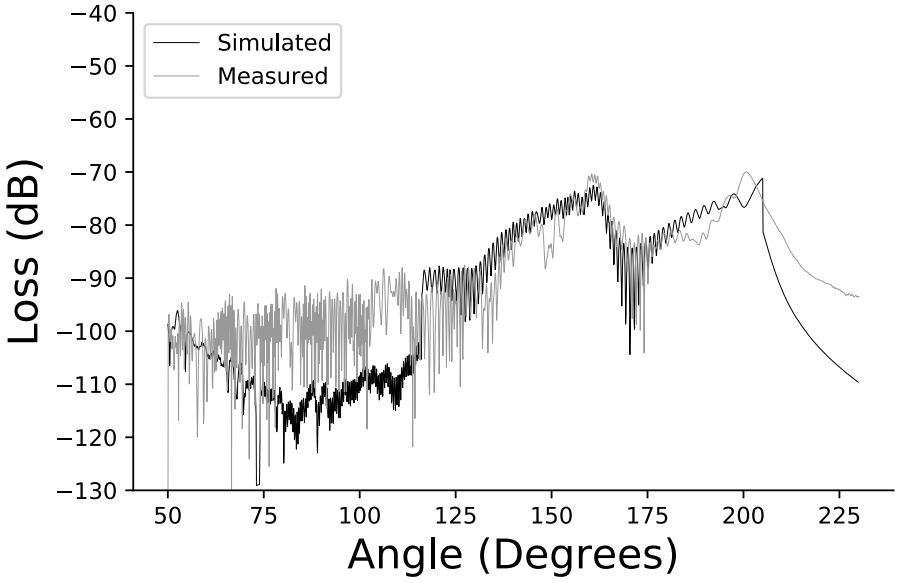


Figure 5.17: Comparison of measurements and simulations for the smooth aluminium coated wedge.

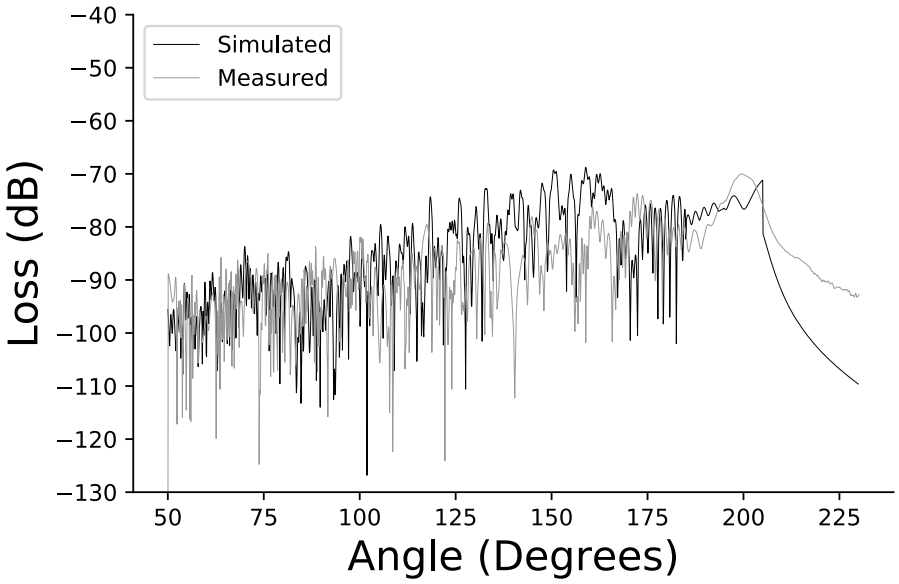


Figure 5.18: Comparison of simulated and measured data for the rough wedge with the standard diffraction coefficients.

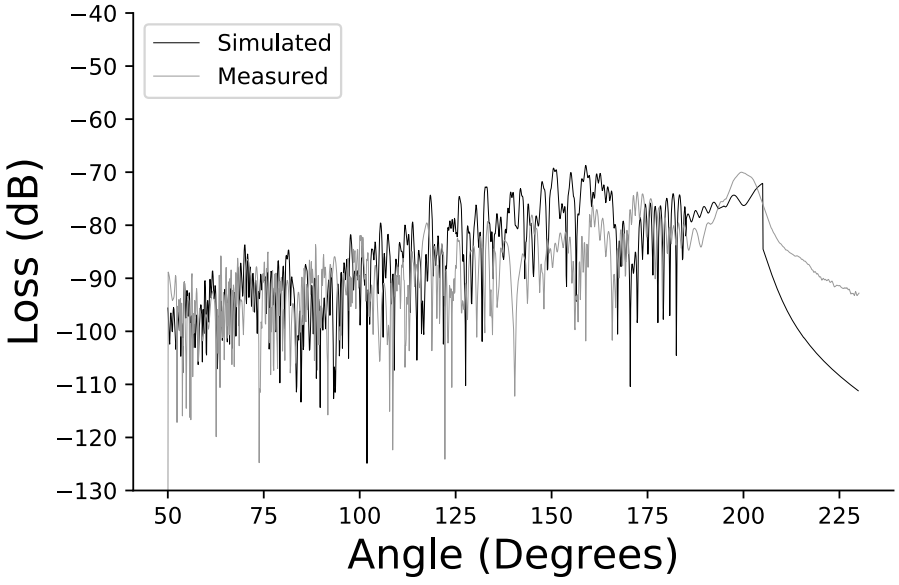


Figure 5.19: Comparison of simulated and measured data for the rough wedge with diffraction coefficients reduced by Rayleigh roughness correction factor.

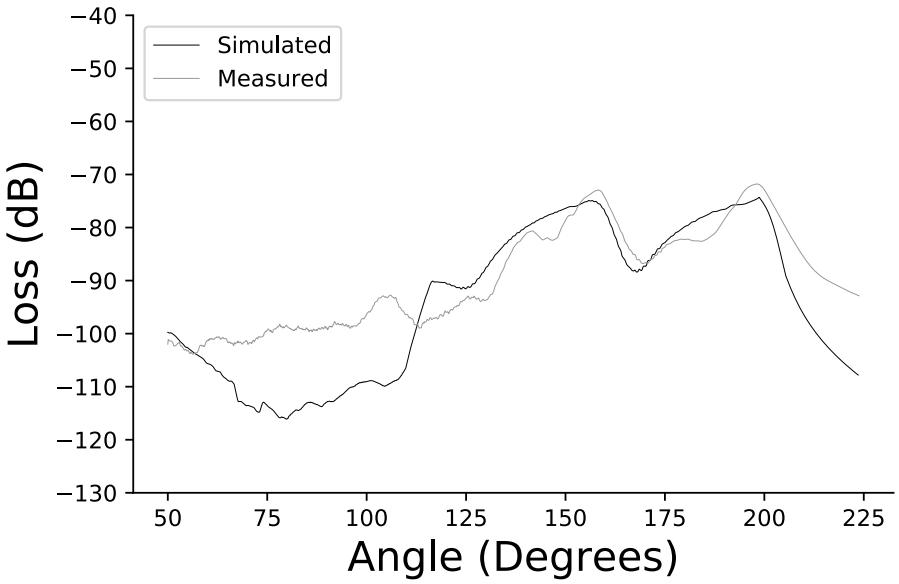


Figure 5.20: Sliding average filtered results for the smooth wedge.

5.8 Conclusion

In this chapter diffuse scattering and diffraction have been investigated as propagation phenomena. Controlled experiments have allowed computer implementations of numerical models to be compared to measurement data in which there is limited opportunity for spurious signals to interfere with the propagation mechanism being investigated.

5.8.1 Scattering

It has been shown that the Kirchhoff model for diffuse scattering reproduces scattering behaviour reasonably well. Unfortunately, the model does not take into account statistical variation across the surface and this might explain discrepancies between measured data and simulated data.

5.8.2 Diffraction

The uniform theory of diffraction was also investigated for mmWave systems. Despite it being widely reported that mmWave frequencies do not diffract efficiently, measured results show that more power was diffracted than the theoretical models suggest for a conductive wedge. The comparisons also demonstrated some of the limitations of ray based models where particular ray paths are either 'on' or 'off', some discontinuities in received power in simulated data were not replicated in measurements. In the unlit region the tail was significantly less than measured results. This means when UTD is used the diffracted power is likely to be lesser than in the real world.

Bibliography

[1] Faroarm - portable 3d measurement arms for any application. https://www.faro.com/en-gb/products/3d-manufacturing/faroarm/.

[2] Products. https://www.anokiwave.com/products/index.html.

[3] W. S. Ament. Toward a theory of reflection by a rough surface. *Proceedings of the IRE*, 41(1):142–146, Jan 1953.

[4] Isaac Amundson, Janos Sallai, Xenofon Koutsoukos, and Akos Ledeczi. Radio interferometric angle of arrival estimation. In Jorge Sá Silva, Bhaskar Krishnamachari, and Fernando Boavida, editors, *Wireless Sensor Networks*, pages 1–16, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.

[5] R. Arablouei and A. Ghorbani. A new utd-based model for multiple diffractions by buildings. In *2002 3rd International Conference on Microwave and Millimeter Wave Technology, 2002. Proceedings. ICMMT 2002.*, pages 484–488, Aug 2002.

[6] Leyre Azpilicueta, Erik Aguirre, Peio López-Iturri, and Francisco Falcone. An accurate utd extension to a ray-launching algorithm for the analysis of complex indoor radio environments. *Journal of Electromagnetic Waves and Applications*, 30(1):43–60, 2016.

[7] T. H. Barratt, E. Mellios, P. Cain, A. R. Nix, and M. A. Beach. Measured and modelled corner diffraction at millimetre wave frequencies. In *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pages 1–5, Sept 2016.

[8] Petr Beckmann and Andre Spizzichino. *The Scattering of Electromagnetic Waves from Rough Surfaces (Artech House Radar Library)*. Artech Print on Demand, 1987.

[9] V. Degli-Esposti. A diffuse scattering model for urban propagation prediction. *IEEE Transactions on Antennas and Propagation*, 49(7):1111–1113, Jul 2001.

[10] V. Degli-Esposti. A diffuse scattering model for urban propagation prediction. *IEEE Transactions on Antennas and Propagation*, 49(7):1111–1113, July 2001.

[11] V. Degli-Esposti, J. S. Lu, J. N. Wu, J. J. Zhu, J. A. Blaha, E. M. Vitucci, F. Fuschini, and M. Barbiroli. A semi-deterministic model for outdoor-to-indoor prediction in urban areas. *IEEE Antennas and Wireless Propagation Letters*, 16:2412–2415, 2017.

[12] Vittorio Degli-Esposti, Franco Fuschini, Enrico Vitucci, Marina Barbiroli, Marco Zoli, Li Tian, Xuefeng Yin, Diego Dupleich, Robert Müller, Christian Schneider, and Reiner Thomä. Ray-tracing-based mm-wave beamforming assessment. *IEEE ACCESS*, 2:1314–1325, 10 2014.

[13] B. Duvenhage, K. Bouatouch, and D. G. Kourie. Numerical verification of bidirectional reflectance distribution functions for physical plausibility. In *Proceedings of the South African Institute for Computer Scientists and Information Technologists Conference*, SAICSIT '13, pages 200–208, New York, NY, USA, 2013. ACM.

[14] R. B. Dybdal. Radar cross section measurements. *Proceedings of the IEEE*, 75(4):498–516, April 1987.

[15] Vuokko Nurmelay et al. Deliverable d1.4 metis channel models. 2 2015.

[16] F. Fuschini, S. Häfner, M. Zoli, R. Müller, E. M. Vitucci, D. Dupleich, M. Barbiroli, J. Luo, E. Schulz, V. Degli-Esposti, and R. S. Thomä. Analysis of in-room mm-wave propagation: Directional channel measurements and ray tracing simulations. *Journal of Infrared, Millimeter, and Terahertz Waves*, 38(6):727–744, Jun 2017.

[17] Abhijeet Ghosh. *Cook-Torrance Model*, pages 146–152. Springer US, Boston, MA, 2014.

[18] A. A. Goulianos, A. L. Freire, T. Barratt, E. Mellios, P. Cain, M. Rumney, A. Nix, and M. Beach. Measurements and characterisation of surface scattering at 60 ghz. In *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, pages 1–5, Sep. 2017.

[19] A. Hsiao, C. Yang, T. Wang, I. Lin, and W. Liao. Ray tracing simulations for millimeter wave propagation in 5g wireless communications. In *2017 IEEE International Symposium on Antennas and Propagation USNC/URSI National Radio Science Meeting*, pages 1901–1902, July 2017.

[20] M. Inomata, T. Imai, K. Kitao, Y. Okumura, M. Sasaki, and Y. Takatori. Prediction accuracy of hybrid method based on ray-tracing and effective roughness model in indoor environment for millimeter waves. In *2017 IEEE Conference on Antenna Measurements Applications (CAMA)*, pages 44–46, Dec 2017.

[21] C. Jansen, S. Priebe, C. Moller, M. Jacob, H. Dierke, M. Koch, and T. Kürner. Diffuse scattering from rough surfaces in thz communication channels. *IEEE Transactions on Terahertz Science and Technology*, 1(2):462–472, Nov 2011.

[22] J. Järveläinen and K. Haneda. Sixty gigahertz indoor radio wave propagation prediction method based on full scattering model. *Radio Science*, 49(4):293–305, 2014.

[23] Joseph B. Keller. Geometrical theory of diffraction*. *J. Opt. Soc. Am.*, 52(2):116–130, Feb 1962.

[24] R. G. Kouyoumjian and P. H. Pathak. A uniform geometrical theory of diffraction for an edge in a perfectly conducting surface. *Proceedings of the IEEE*, 62(11):1448–1461, Nov 1974.

[25] J. Lee, J. Choi, J. Lee, and S. Kim. Permittivity effect of building materials on 28 ghz mmwave channel using 3d ray tracing simulation. In *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, pages 1–6, Dec 2017.

[26] R. Luebbers. Propagation prediction for hilly terrain using gtd wedge diffraction. *IEEE Transactions on Antennas and Propagation*, 32(9):951–955, Sep. 1984.

[27] B. Neekzad, K. Sayrafian-Pour, J. Perez, and J. S. Baras. Comparison of ray tracing simulations and millimeter wave channel sounding measurements. In *2007 IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications*, pages 1–5, Sep. 2007.

[28] K.H. Ng, E.K. Tameh, and A.R. Nix. A new heuristic geometrical approach for finding non-coplanar multiple edge diffraction ray paths. *Antennas and Propagation, IEEE Transactions on*, 54:2669 – 2672, 10 2006.

[29] Michael Oren and Shree K. Nayar. Generalization of the lambertian model and implications for machine vision. *International Journal of Computer Vision*, 14(3):227–251, Apr 1995.

[30] J. Pascual-García, J. M. Molina-García-Pardo, M. T. Martínez-Inglés, J. V. Rodríguez, and N. Saurín-Serrano. On the importance of diffuse scattering model parameterization in indoor wireless channels at mm-wave frequencies. *IEEE Access*, 4:688–701, 2016.

[31] Bui Tuong Phong. Illumination for computer generated pictures. *Commun. ACM*, 18(6):311–317, June 1975.

[32] Bui Tuong Phong. Illumination for computer generated pictures. *Commun. ACM*, 18(6):311–317, June 1975.

[33] R. Piesiewicz, C. Jansen, D. Mittleman, T. Kleine-Ostmann, M. Koch, and T. Kurner. Scattering analysis for the modeling of thz communication systems. *IEEE Transactions on Antennas and Propagation*, 55(11):3002–3009, Nov 2007.

[34] S. Priebe, M. Jacob, C. Jansen, and T. Kürner. Non-specular scattering modeling for thz propagation simulations. In *Proceedings of the 5th European Conference on Antennas and Propagation (EUCAP)*, pages 1–5, April 2011.

[35] S. Priebe, M. Jacob, C. Jansen, and T. Kürner. Non-specular scattering modeling for thz propagation simulations. In *Proceedings of the 5th European Conference on Antennas and Propagation (EUCAP)*, pages 1–5, April 2011.

[36] S. Priebe, M. Jacob, and T. Kürner. Polarization investigation of rough surface scattering for thz propagation modeling. In *Proceedings of the 5th European Conference on Antennas and Propagation (EUCAP)*, pages 24–28, April 2011.

[37] S. Priebe, M. Jacob, and T. Kürner. Polarization investigation of rough surface scattering for thz propagation modeling. In *Proceedings of the 5th European Conference on Antennas and Propagation (EUCAP)*, pages 24–28, April 2011.

[38] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez. Millimeter wave mobile communications for 5g cellular: It will work! *IEEE Access*, 1:335–349, 2013.

[39] Remcom. *5G mmWave Channel Modeling with Diffuse Scattering in an Office Environment. ebook.* (n.d.).

[40] K. A. Remley, H. R. Anderson, and A. Weisshar. Improving the accuracy of ray-tracing techniques for indoor propagation modeling. *IEEE Transactions on Vehicular Technology*, 49(6):2350–2358, Nov 2000.

[41] Karim Rizk, Reinaldo Valenzuela, Steven Fortune, D Chizhik, and F Gardiol. Lateral, full-3d and vertical plane propagation in microcells and small cells. volume 2, pages 998 – 1003 vol.2, 06 1998.

[42] J M Saiz, F Gonzalez, F Moreno, and P J Valle. Application of a ray-tracing model to the study of back scattering from surfaces with particles. *Journal of Physics D: Applied Physics*, 28(6):1040–1046, jun 1995.

[43] Simon R. Saunders. *Antennas and Propagation for Wireless Communication Systems*. Wiley, sep 1999.

[44] Freire-AL Sayer, L. mmwave propagation around rough and smooth conducting wedges. In *The 2020 IEEE 91st Vehicular Technology Conference (Submitted)*, 2020.

[45] Lawrence Sayer, Alberto Loaiza Freire, Evangelos Mellios, and Andrew Nix. A kirchhoff scattering model for millimetre wavelength wireless links. In *2018 12th European Conference on Antennas and Propagation (EUCAP 2018)*, 2018.

[46] K. R. Schaubach, N. J. Davis, and T. S. Rappaport. A ray tracing method for predicting path loss and delay spread in microcellular environments. In *[1992 Proceedings] Vehicular Technology Society 42nd VTS Conference - Frontiers of Technology*, pages 932–935 vol.2, May 1992.

[47] D. Solomitckii, Q. C. Li, T. Balercia, C. R. C. M. da Silva, S. Talwar, S. Andreev, and Y. Koucheryavy. Characterizing the impact of diffuse scattering in urban millimeter-wave deployments. *IEEE Wireless Communications Letters*, 5(4):432–435, Aug 2016.

[48] Dmitrii Solomitckii, Qian Clara Li, Tommaso Balercia, Claudio R C M Da Silva, Shilpa Talwar, Sergey Andreev, and Yevgeni Koucheryavy. Characterizing the impact of diffuse scattering in urban millimeter-wave deployments. *IEEE Wireless Communications Letters*, 5(4):432–435, 8 2016.

[49] S. Ahsanuzzaman Md Tariq, C. Despins, S. Affes, and C. Nerguizian. Rough surface scattering analysis at 60 ghz in an underground mine gallery. In *2014 IEEE International Conference on Communications Workshops (ICC)*, pages 724–729, June 2014.

[50] Agilent Technologies. Basics of measuring the dielectric properties of materials. `http://cp.literature.agilent.com/litweb/pdf/5989-2589EN.pdf`, 2000. [Online; accessed 14th August 2017].

[51] A. Torabi and Amir Ahmad Shishegar. Scattering computation of slightly rough gaussian surfaces by kirchhoff approach with harmonic extraction analysis. In *2009 International Conference on Microwave Technology and Computational Electromagnetics (ICMTCE 2009)*, pages 292–295, Nov 2009.

[52] K. E. Torrance and E. M. Sparrow. Theory for off-specular reflection from roughened surfaces*. *J. Opt. Soc. Am.*, *57*(9):1105–1114, Sep 1967.

[53] E. M. Vitucci, F. Mani, V. Degli-Esposti, and C. Oestges. Polarimetric properties of diffuse scattering from building walls: Experimental parameterization of a ray-tracing model. *IEEE Transactions on Antennas and Propagation*, 60(6):2961–2969, June 2012.

[54] Bruce Walter, Stephen R. Marschner, Hongsong Li, and Kenneth E. Torrance. Microfacet models for refraction through rough surfaces. In *Proceedings of the 18th Eurographics Conference on Rendering Techniques*, EGSR'07, pages 195–206, Aire-la-Ville, Switzerland, Switzerland, 2007. Eurographics Association.

[55] Stephen Westin, Hongsong Li, and Kenneth E. Torrance. A comparison of four brdf models. 06 2019.

[56] Z. Yun and M. F. Iskander. Ray tracing for radio propagation modeling: Principles and applications. *IEEE Access*, 3:1089–1100, 2015.

Chapter 6

Optimisation of Ray Tracing

6.1 Introduction and Definitions

In this chapter, ray model optimisation techniques are presented. The chapter builds on work more familiar to the computer graphics community. This is because in certain graphics rendering applications the intersection of rays with 3D compute models of an environment is central to timely rendering of computer generated models. In this chapter the term 'primitive' is used. This refers in general to a planar polygon, but in the context of this chapter is used mainly to refer to a triangle.

Pl"ker coordinates are used later in the chapter. A familiar way of defining a line in euclidean 3-space is to specify two points, say p_1 and P_2, each with an x, y and z coordinate. An alternative means of uniquely specifying the line is to store a direction, $D = P_1 - P_2$ and a moment around the origin $M = P_1 \times P_2$ where $\times$ is the vector cross product. Again six numbers are used, but storing the line in this form allows for certain very efficient methods of determining which way a line passes another to be used which allows programs to be faster.

6.2 Spatial Technologies

MIMO and mmWave technologies will be used to achieve the higher data-rates and spectral-efficiency necessary to deliver 5G services [22][14] and beyond. These technologies either exploit the spatial characteristics of the channel, as in the case of MIMO systems, or, in the case of mmWave beam-forming, means that where the signal is 'coming from' is very important. Ray models have seen a resurgence in popularity as an attractive modelling solution [5][1][27][13]. They are provided with a digital representation of an environment and, therefore, the channel extracted from simulations is inherently 'spatial'. The interactions of the emitted electromagnetic wave with the environment are inherently captured. This facilitates a sound understanding of spatial and temporal channel

properties, and hence contributes to more informed network and device design.

To properly design networks that make use of spatial technologies - like beamforming and MIMO - it is desirable to track phase relationships between separate array elements and users. Simply assuming a random uniform phase for each multipath component derived from ray tracing may not give realistic results. Unrealistic channel data would hamper efforts to benchmark algorithms for beam-forming, precoding and other processes. Recent research has focused on the ability to simulate the small scale effects of a channel, not only for a single point in time and space, but across multiple points in time and/or space. This means accounting for the correlation of the channel over small separations in either domain. In [17][16], a method for predicating the phase evolution of the popular effective roughness (ER) model was presented. Works such as [4] and [20] have demonstrated exciting new techniques to retrieve the Doppler shifts in the channel from a "single shot" of ray tracing. This means, within the time that propagation paths remain broadly the same - i.e. they do not drift off the end of a reflecting surface or into a blocker - any resolution of time steps can be simulated using only a single path finding operation. Since path finding is the most intensive operation in any ray tracing software this is a useful optimisation. Unfortunately, for some systems - for example distributed MIMO networks - the channel will vary between disparate 'array' elements significantly, and hence many individual element to element path finding operations must be performed for a very high number of links. The same will be true in other cases, for example, high mobility scenarios where significant propagation paths change frequently.

In this chapter, a ray model has been implemented. The tool is designed to accurately track relative phase of multipath components, correctly accounting for phase differences between disparate paths for many point to point links by using the image method. The tool is used mainly as a platform for presenting the benefits of some novel optimisation strategies that apply to the ray-path finding process. These mitigate some of the computational resource requirements of the ray tracing operation.

6.3 Justification for Choice of Ray Paradigm for Phase Consistent Predictions

It was discussed in Chapter 3 that one categorisation of ray models is based upon the way paths are found between transceiver locations. Models may be divided into ray tracers (image method), and ray launchers (shoot bouncing ray, ray casting, 'pin cushion' method) [9]. In ray launching schemes, rays are fired discretely into the environment and are bounced around according to the rules of reflection, transmission, diffraction and diffuse-scattering, before coming close enough to the RX point to count as received [8]. In ray tracing, imaging is used to construct virtual sources through primitives and, from these virtual sources, straight line paths between terminals and primitives are constructed [2]. Many state of the art tools follow a hybrid approach, mixing methods for different

propagation phenomena, typically using ray launching for diffuse scattering (DS) and ray tracing for other interactions [9].

6.3.1 Computational Effort and Error

Ray launching scales more favourably than ray tracing when the maximum allowed interaction order is increased [6], but introduces path length errors. An analysis of why this is so is presented here.

Ray Launching

When ray launching, for an environment with κ primitives, with maximum allowed reflection order n_{max} and r initially fired rays, the maximum amount of computational effort, c, is given by (6.1).

$$c = rn_{max}\kappa \tag{6.1}$$

Equation 6.1 is true when all rays stay within the environment until the maximum reflection order is reached. In practice state of the art ray launchers will destroy a ray path once a propagation out into space is initiated, or if a path is too long, meaning (6.1) is an upper bound. The number of primitives, κ, is usually fixed for an environment, and n is fixed based on required accuracy of predictions (usually limited to less than 7). However, the choice of total number of rays to fire, r, is somewhat arbitrary. It decides the spatial resolution of the ray launching process. Both 'pin cushion' and 'environment driven' ray launchers introduce errors in path length, and therefore, a predicted phase error. The small errors in path length and/or direction do not affect power predictions significantly [15]. If accurate relative phase information is required - for simulating fast fading, or relative phase across a MIMO array for example - then the angular separation between rays, or spatial discretisation of the environment, must be such that the total error in path length is some fraction of a wavelength. An error of $\frac{1}{8}^{th}\lambda$ is a reasonable suggestion for most purposes since it ensures that two multipath components that should constructively interfere do indeed have a constructive interference with a relatively small field level and phase error. Often, a very high discretisation-resolution must be used. This causes execution to be slow. For a 3.5 GHz system, the spacing between tiles in an environment driven ray launcher would have to be about 1 cm. This means for every square metre of environment, ten thousand rays would need to be fired. Even in a small room this could mean millions of rays, and for a large city this resolution becomes completely impractical. Indeed, this type of problem was illustrated when comparing the results from the highly optimised ray launcher used for the work in [26] and presented in chapter 4 of this thesis - which uses an angular discretisation - to an analytical two ray model as part of [19]. A simulation was performed in which a transmitter broadcast to a a grid of dipoles over a flat earth. The time taken to achieve power prediction accuracy on each path to within 0.1 dB was about ten minutes on

a desktop machine with a 3.4 GHz processor and 16 GB RAM using 8 threads, whereas to achieve one eighth wavelength phase accuracy, it took eight hours.

Ray Tracing

For a naive ray tracing approach, an image source must be created for every primitive in the environment, and for every primitive apart from the one just used to generate the previous image from then on. Once each image source is generated, the path must be checked against every other primitive for obstruction for the n ray segments in that path. This leads to an algorithmic complexity given by

$$c = \sum_{n=1}^{n_{max}} \left[\kappa . \left[(\kappa - 1)^{(n-1)} \right] \times n\kappa \right] \tag{6.2}$$

Again this is an upper bound because, for example, when checking the κ intersections per path segment, if the reflecting primitive is missed, or at least one blocker is hit early on, subsequent primitives and path segments need not be checked. This expression is exponential in the total number of allowed reflections, and where k is large, as is typical for large environments, this can cause calculation of valid paths to be prohibitively slow. However, ray tracing has no discretisation error. This means for phase consistent predictions, it could in fact be faster than ray launching if proper optimisation was performed. The focus of this chapter is the implementation of a ray tracing tool with some novel optimisation techniques applied.

6.4 Ray Model Optimisation Techniques Review

Since ray tracing is a reasonably complex operation, many methods of optimisation have been invented, and a review of these is provided here to contextualise the content of this chapter.

6.4.1 Efficient Intersection Tests

A heavily repeated operation in all ray modelling is the intersection of a ray with primitives in the environment. As such, this operation is critical to the overall execution speed of the program. Several efficient tests have been proposed. If minimum storage is a requirement, for memory constrained applications, then the Möller Trombore algorithm is widely used [18]. If a mesh can be prepossessed and intermediate results stored, a method based on Plücker coordinates requires less operations per intersection test.

6.4.2 Parallelism

In ray modelling, a few elementary operations are repeated many times. An example would be the intersection of a ray with the primitives in the environment. The result of each intersection test does not rely on the result of the others, and this means this is an ideal candidate for parallelism.

The use of GPUs enables the ray tracing operation to be performed extremely quickly [10]. There are many other stages at which parallelism may be applied, indeed in the ray launching tool described in chapter 4, separate threads are created for each initial new ray and the tree structure is traversed in parallel on multiple CPUs.

6.4.3 Elliptical Exclusion Zone

Ray paths that travel further from transceiver locations will generally be weaker than those who have shorter propagation routes due to path loss. If it is desired to simulate the channel between two transceiver locations in a model of a whole city, it is likely that only those structures closest to the transceiver locations will actually contribute significantly to the simulation. For this reason, including far away structures in simulations is an unnecessary computational burden. A method for removing parts of the mesh that are extraneous is drawing an ellipse around the two transceiver locations. Only parts of a mesh that lie within this ellipse need to be counted in simulations. This may be done very efficiently given the location of a particular pair, as in [3].

6.4.4 BVH or Octree Decomposition

BVHs and Octrees [7] are two similar methods used to reduce the number of intersection tests that must be performed to find a ray primitive intersection. Both work by subdividing the primitives by spatial location into bounding boxes. A simple ray to bounding-box intersection test may be performed and only if a hit is found do the primitives inside need to be checked for intersection with a ray. By nesting such bounding boxes the final κ terms in both 6.1 and 6.2 may be reduced to $\log(\kappa)$, significantly reducing computation time.

6.4.5 Use of a Visibility Matrix

If a scene is fixed, pre-computing a visibility matrix allows for the reduction in the size of the path search tree in ray tracing, and reduces the number of intersection tests that must be performed for both ray tracing and ray launching. The intra-mesh visibility matrix is a stored list of visibility relations between primitives in a 3D model of an environment. It is generally referred to as a 'from-region potentially-visible set' by the computer graphics community. A more detailed discussion of the idea can be found in [9][12][11]. In previous literature from the propagation modelling community, it is constructed in a non robust way, either by just checking the path between the centres of primitives for blockage, as in [15], or by trying lots of possible paths between different locations on pairs of primitives as in [25]. The more sophisticated methods presented in [6] and [11] apply to 2D environments, and cannot be directly extended into three dimensions.

Fig 6.1 shows how a visibility matrix may be used to construct virtual image sources, and therefore, find paths between transmitter and receiver in an efficient way [6]. From the transmitting

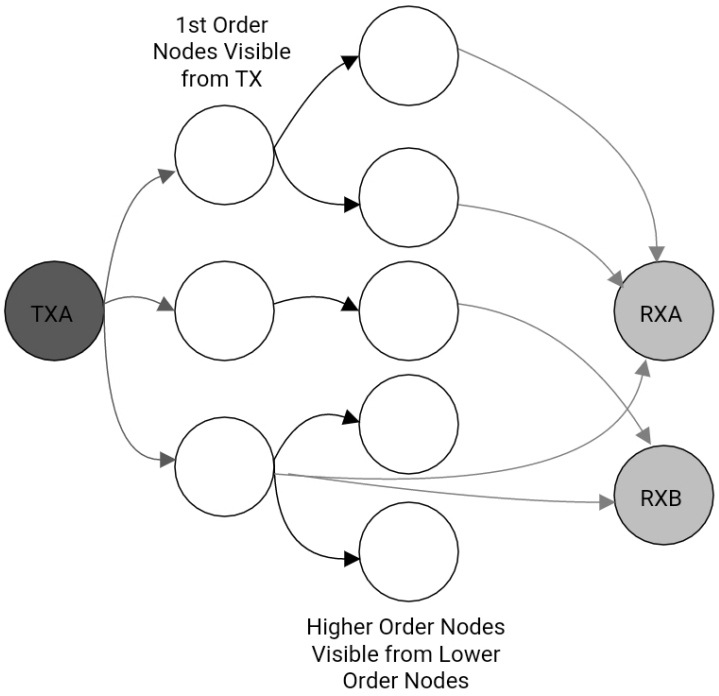


Figure 6.1: Graph representation of path finding operation.

node, all primitives directly visible from the transmitter are connected as first layer nodes. If any of these nodes are visible from an RX node the associated path is validated. Then, the visibility matrix is used to generate second layer nodes. These nodes are those that are visible - or at least partially visible - from the first layer nodes. Again for this layer, any nodes visible to an RX node have an associated path validation step. Layers of nodes are added until a maximum interaction order is reached. When considering only reflections, and a specific TX location to generate image sources, at each level some of the possibly visible nodes may be pruned, because they are not actually visible from the given image source after a previous reflection. The visibility matrix can help with this pruning step too, because fewer primitives need to be eliminated for a specific image source location. Paths are validated by checking that the path generated through the particular secession of primitives actually intersects those primitives, and also is not blocked by another primitive not in the ordered set. This basic method, shown in Fig 6.1, was presented in [6]. A difference in the method in this work is that there are many RX nodes in the graph. This avoids repeated graph construction for every RX location (it is likely that this was done in [6] but not made explicit). In this work, a preprocessing step is performed to speed up the finding of the initial links from the TX and the final links to the RX. This was achieved by first finding the visibility of the mesh from a large plane, say 1 m from the ground. If transceiver locations are constrained to this plane, then it was possible to exclude any primitives that are not visible from the plane - considered as a whole - from consideration as visible to a particular transceiver point. If there are many transceiver points this

can be a significant time saving mechanism. For the purposes of this work one plane was generated at roof height for the BS elements, and one 1 m from the ground for UI locations.

6.5 Calculation of the Intra-mesh Visibility Matrix

6.5.1 Background

In [6], a method was presented in which the visibility or invisibility between two straight line primitives under a set of blockers could be represented as a 2D polygon. This method enabled any mutually invisible primitives to be identified as such. Furthermore, for partially visible primitives at the path validation stage, a very efficient point in polygon algorithm was used to verify if a specific path between them was blocked or not. Unfortunately when attempting to extend this idea directly into thee dimensions, constructive solid geometry must be performed between four-dimensional solids with non-planar surfaces. This means there is no direct equivalent to the simple "point in polygon" check used in [6], and that it is more difficult to store the visibility relations between primitives. However, some ideas may be extended into three dimensions, and in this work an aggressive method to calculate the potential visibility set from each primitive is presented. The method arbitrarily approaches an exact solution to intra-mesh visibility.

6.5.2 First Pass Visibility

For finding the visibility of a mesh from a point, backface culling may be used to remove invisible faces [6]. This cannot be used for primitive to primitive visibility, because a primitive might be backwards facing for some locations on the other primitive and front facing for others. However, a triangle fully behind another triangle cannot be visible from the triangle that it is behind. Polygons with identical normals may be determined to be mutually invisible, this extending from the fact that all polygons in the environment are part of a solid of finite thickness, so two polygons facing the same direction will necessarily be blocked by the back-side of one of them. The visibility from a triangular primitive, that is marked in blue, is shown for a simple scene after this first pass visibility sweep in Fig 6.2.

6.5.3 Intra-mesh Visibility Method Overview

Consider a 3D mesh entirely comprised of triangular primitives. Between any pair of triangles, there are potentially many blocking triangles that partially or fully obscure visibility between the pair of interest. Any location on a primitive can be described by two coordinates in a 2D coordinate system on the plane of the triangular primitive. This means two locations - one each on on two triangles - can be specified with four variable instances. Each blocking primitive as well as the arrangement of the two primitives of interest can be codified as a series of constraint inequalities on

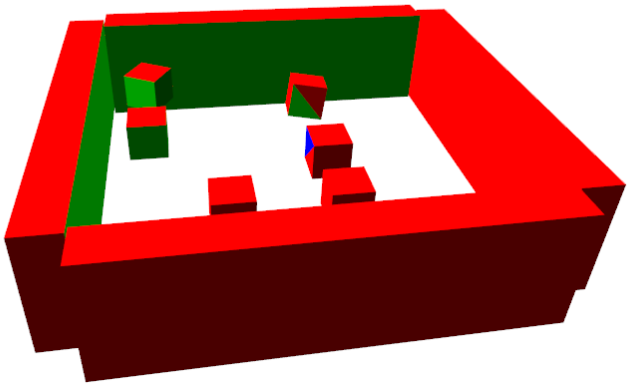


Figure 6.2: A simple scene showing triangles visible from the blue triangle in green, and invisible triangles in red, after a first pass visibility sweep.

these four variables. To determine if a path is valid between points on two triangles, the constraint equations must be evaluated for the four instances specifying the specific endpoints. More generally, to determine if any path exists between two primitives that is unblocked, a feasible region under the set of constraints must be searched for in four dimensional space. If one or more exists then there is at least one straight line path between the two primitives, and they must be counted as at least partially mutually-visible. The method for creating the constraint equations relies upon 3D geometry, and is presented in the next section. Once the constraint equations have been formed, a feasible region is searched for by use of the mixed integer non-linear programming library MIDACO [23]. Depending upon the maximum number of search iterations that MIDACO is allowed to perform, the algorithm may arbitrarily approach an exact visibility solution. In practice, about 2000 iterations of the search algorithm is all that is needed before convergence is reached in practice. Having used the constraints to find if two primitives are visible at all, they may be re-used later at the path validation stage for a path between two particular locations on a pair of triangles. In this way the coefficients for the constraint inequalities may be thought of in a similar way to the polygons representing the visibility relationships in [6].

6.5.4 Determining Potential Blockers

Naively, all triangles could be considered as blockers between any pair of triangles in a scene, apart from the paired triangles themselves. For a scene with κ primitives, there are $\frac{\kappa^2-\kappa}{2}$ ordered pairs of different - i.e. a triangle may not form a valid pair with itself - triangles. For each ordered pair $\kappa - 2$ blockers would be considered.

If some constraints are stored per blocker per pair, even for relatively small mesh, memory re-

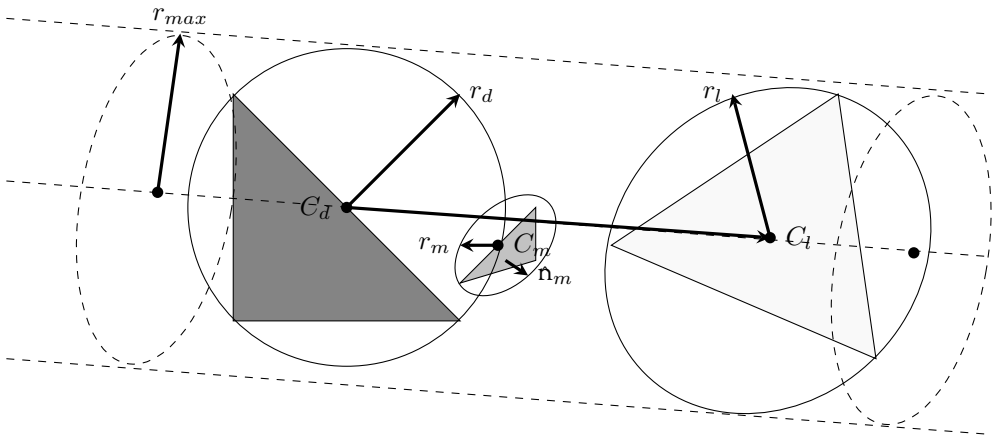


Figure 6.3: Blocker being identified between a pair of triangles.

quirements become indefeasibly large. A simple way of reducing required memory is by reducing the $\kappa - 2$ blockers that must be stored for each ordered pair. It is important to not store any blockers that are not 'between' the pair. To be between a pair, a blocker must have survived the first pass visibility check for the source triangle. Likewise if a blocker is in a lateral direction outside the smallest area bounding the source/destination pair it cannot possibly block a path between the pair.

Whilst finding the smallest bounding volume of the triangles making a pair, and then performing a blocking triangle intersection test with this, perhaps complex, shape is optimal for reducing false positives, in practice it is difficult to implement and more importantly slow to execute. A compromise between speed and accuracy is presented here. First the circumcentre of each triangle is found. The definition of the circumcentre is the point on the plane of the triangle such that the distance from the point to each vertex is the same. The circumcentres are denoted as C in Fig 6.3. The circumcentre is found using the following formula for a triangle with vertices A, B and D [24],

$$C = \frac{1}{2|\vec{AB} \times \vec{AB}|^2}\left[[(|\vec{AD}|^2(\vec{AB} \times \vec{AD})) \times \vec{AB}] + [(|\vec{AB}|^2\vec{AD}) \times (\vec{AB} \times \vec{AD})]\right] + A. \tag{6.3}$$

To determine if a blocking triangle is "between" a pair of triangles, its circumcentre - shown as C_m in Fig 6.3 - radius, r_m, and normal, $\hat{\mathfrak{n}}_m$ are used. The vector between the circumcentres of the pair of triangles that it potentially blocks is constructed - $\vec{C_dC_l}$ in Fig 6.3. The radius r_{max}, which is the greater of the two radii of the circumscribing circles of the source and target triangle is also retrieved. If the perpendicular distance between the circumcentre of the blocking triangle and the line between the circumcentres of our pair of triangles is δ, then our triangle is "between" the pair iff

$$\delta < r_{max} + r_m \times abs\left(\frac{\hat{\mathfrak{n}}_m \cdot \vec{C_dC_l}}{|\vec{C_dC_l}|}\right). \tag{6.4}$$

Equation (6.4) does not take into account the fact that the closest point of the blocking triangle to

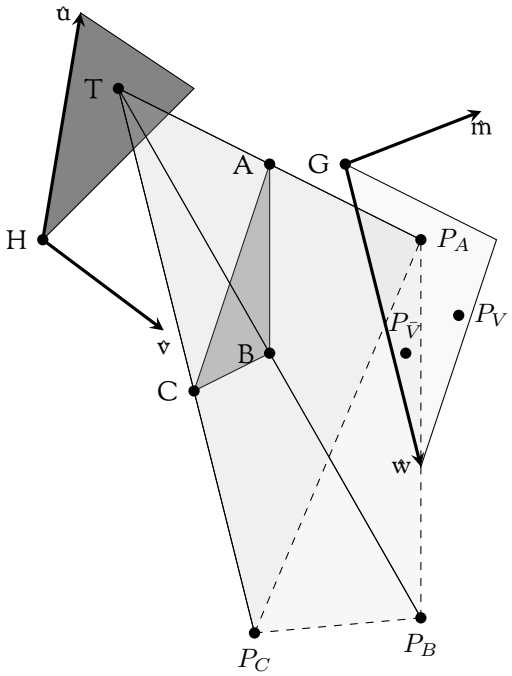


Figure 6.4: Visibility of the light grey triangle from the dark grey triangle blocked by medium grey triangle.

the line that travels through C_b and C_y, might be behind one of the triangles in the pair. The first answer to this apparent problem is that triangles that are *fully* behind either triangle are removed from this check altogether. It is more advantageous to use two quicker "fully behind" checks (the result can be remembered for every ordered pair) than one of the checks described here. The other case, that may arise, is that the closest point on a particular triangle to the line is behind one of the triangles, and by the time it emerges in front of the aforementioned triangle, it is much further away from the line and therefore too far to be a blocker. This will indeed generate a false positive, but in practice, the extra computational cost involved with handling these cases more correctly is anticipated to be more detrimental than the advantage of discounting such blockers in most real geometries.

By finding what primitives lie between any pair in the mesh, the number of blockers considered per pair may be radically reduced to a small fraction of the total number of primitives in the mesh. Finding blockers between a transceiver point and a triangle proceeds in the same way, except the radius of a point is zero, and hence it is the radius of the (only) triangle that is used as r_{max} in (6.4).

6.5.5 Visibility Under Blockage

Introduction

A general blocking situation is shown in Fig 6.4. All triangles are arranged arbitrarily in 3D space. The objective is to compute the visibility of the light grey triangle from the dark grey triangle in the presence of blocking triangles.

In Fig 6.4, the vectors $\hat{\mathbf{u}}$ on the dark grey triangle and $\mathbf{w}$ on the light grey triangle are shown. These are the vectors along the longest edge of each triangle. The vectors $\hat{\mathbf{v}}$ and $\hat{\mathbf{m}}$ are perpendicular to $\hat{\mathbf{u}}$ and $\mathbf{w}$ respectively in the plane of their respective triangles. Using the longest edge as the base of the triangle, the magnitude of $\hat{\mathbf{v}}$ and $\hat{\mathbf{m}}$ is equal to the height of their respective triangles. In the local coordinate systems, any point on the dark grey triangle may be specified as (p, q) and on the light grey triangle as (s, t). To convert these into a global point in the three dimensional coordinate system with H as the origin of the dark grey coordinate system and G as the origin of the light grey's coordinate system:

$$P_d = H + p\hat{\mathbf{u}} + q\hat{\mathbf{v}}, \tag{6.5}$$

$$P_l = G + s\mathbf{w} + t\hat{\mathbf{m}}. \tag{6.6}$$

There are three types of constraint on p, q, s and t that act to determine visibility. The first type of constraint (Type 1) encodes the shadowing caused by a blocking triangle. The second type is imposed by the relative positions of all the triangles (Type 2). The third kind of constraint encodes that a point must actually be in a triangle. Since the coordinates can specify any point on the infinite plane containing the triangle, constraints must be used to restrict us to locations within the triangle (Type 3).

Type 1 Constraints

In Fig 6.4 a blocking triangle is shown in medium grey. Observing from the point T, the invisible region of the plane of the light grey triangle is triangle (P_A, P_B, P_C). In general this region may not be triangular, and depending on the orientation and position of the three triangles may be infinite in extent. It is found, without loss of generality, by finding the three intersection lines between the plane of the light grey triangle and each of the planes containing T and two of the blocking vertices (i.e. two of A, B or C). These planes are shown in Fig 6.4. Having constructed these lines, an invisible point (e.g. $P_{\bar{V}}$) lies inside all of them whilst a visible point (e.g. P_V) must lie outside at least one of them. In the case of P_V in Fig 6.4, it lies outside the line P_AP_B. "Inside" and "outside" must be determined carefully paying attention to the winding order of T and the two vertices used to create the plane.

In Fig 6.4, the light grey triangle is indeed visible from the dark grey, because there is at least one point on the light grey triangle visible from at least one point on the dark grey triangle, e.g. P_V from

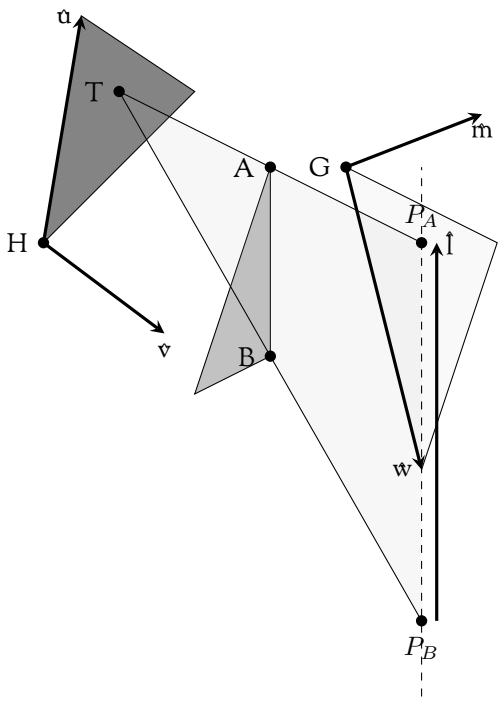


Figure 6.5: Visibility of the light grey triangle from the dark grey triangle blocked by the medium grey triangle.

T. In other cases, for example if the medium grey triangle was much larger, the light grey triangle could be completely obscured. There may also be many blockers creating compound shadows.

The analysis for the formation of the constraint for each edge of a blocking triangle proceeds in the same way, and, for brevity, the plane containing T, A and B is considered here. It will be helpful in the following analysis to refer to Fig 6.5 which removes extraneous details from Fig 6.4.

The constraints are in the form of a series of coefficients for non-linear combinations of the four variables p, q, s and t. Intuition for the mathematical process is given here.

Intuition First T is calculated using (6.5). Then direction vectors $\hat{\mathfrak{a}} = A - T$ and $\hat{\mathfrak{b}} = B - T$ are calculated. Taking the cross product gives the normal of the plane $\hat{\mathfrak{n}}_1$, being careful to choose the correct winding order to get an "outward" facing normal, $\hat{\mathfrak{n}}_1 = \hat{\mathfrak{a}} \times \hat{\mathfrak{b}}$. A point on this plane is identified. in this case A is convenient. A direction vector, $\hat{\mathfrak{l}}$, for the line of intersection between the two planes (i.e. the plane that has just been calculated and the light grey plane) is given by

$$\hat{\mathfrak{l}} = \hat{\mathfrak{n}}_1 \times \hat{\mathfrak{n}}_2, \tag{6.7}$$

where $\hat{\mathfrak{n}}_2$ is the normal of the light grey plane. A point on the line E_1 must be found to fully define it. Note that

$$E_1 = \frac{(d_2\hat{\mathfrak{n}}_1 - d_1\hat{\mathfrak{n}}_2) \times \hat{\mathfrak{l}}}{\hat{\mathfrak{l}} \cdot \hat{\mathfrak{l}}}, \tag{6.8}$$

where

$$d_1 = -A \cdot \hat{\mathfrak{n}}_1 \tag{6.9}$$

and d_2 is calculated in the same way and pre-stored for the light grey triangle, using a point on, and the normal of, the triangle. Using the direction and point on the line, a second point on the line is generated.

$$E_2 = E_1 + \hat{\mathfrak{l}}. \tag{6.10}$$

Note that it appears that P_A and P_B could be used as the two points, however if the arrangement of triangles is such that $T + \rho\hat{\mathfrak{a}}$ and/or $T + \rho\hat{\mathfrak{b}}$, where ρ is a scalar, do not intersect the plane of the light grey triangle this is impossible because P_A and/or P_B do not exist. To determine the equivalent of the points E_1 and E_2 in the local coordinate system of the light grey triangle, D_1 and D_2, it is noted that

$$D_n = \begin{bmatrix} s_n \\ t_n \\ \alpha_n \end{bmatrix} \tag{6.11}$$

where n takes the value 0 or 1, denoting which of the two points is of interest, and where α_n is a dummy variable that should always be equal to zero. This is because E_n is always in the plane of the light grey triangle. This means α_n is ignored when the solution for D_n is found, leaving 2D coordinates (s_n, t_n). The matrix C is constructed by stacking the column vectors $\hat{\mathfrak{w}}$, $\hat{\mathfrak{m}}$ and $\hat{\mathfrak{n}}_2$, in that order, giving

$$\mathrm{C} = \begin{bmatrix} w_x & m_x & n_{2x} \\ w_y & m_y & n_{2y} \\ w_z & m_z & n_{2z} \end{bmatrix}. \tag{6.12}$$

The equation

$$\mathrm{C}D_n = E_n - G. \tag{6.13}$$

must be solved. Rearranging this gives

$$\mathrm{C}^{-1}\mathrm{C}D_n = \mathrm{C}^{-1}(E_n - G), \tag{6.14}$$

and therefore

$$D_n = \mathrm{C}^{-1}(E_n - G). \tag{6.15}$$

This is a means of calculating D_1 and D_2. Fig 6.6 shows D_1 and D_2 comprised of coordinates (s_1, t_1) and (s_2, t_2) respectively. For a point (s, t) to be blocked then the constraint

$$(s - s_1)(t_2 - t_1) - (t - t_1)(s_2 - s_1) < 0, \tag{6.16}$$

must be satisfied. This is equivalent to

$$s(t_2 - t_1) - t(s_2 - s_1) + (t_1 s_2 - t_2 s_1) < 0. \tag{6.17}$$

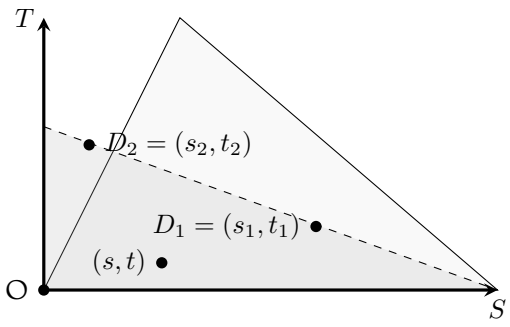


Figure 6.6: Calculating the visibility or invisibility of a point on the light grey triangle

Following the mathematics presented above, but avoiding evaluating any intermediate terms for a specific (p, q) to keep the inequality (6.17) general to any (p, q, s, t), means the resultant inequality consists of 60 coefficients in non linear combinations of p, q, s and t. These coefficients must be calculated per edge per blocker per pair.

Type 2 Constraints

A potential blocking triangle may be behind one of the triangles, in which case it is removed from consideration as a blocker. Likewise, the two triangles, between which visibility is being established, may be one behind the other, and if so would be considered mutually invisible. They may also be fully in front of one another in which case no constraint is needed. However, the planes containing the primitives may intersect one another somewhere in the region of one of the triangles. In this partially overlapping case, constraints are formed.

This problem is illustrated in Fig 6.7, where the medium grey triangle is not in a position to block a path between the dark grey and light grey triangles, but some points on the light grey triangle are not visible from the dark grey triangle, without the path going back through the dark grey triangle, which is impossible if the dark grey triangle is part of a solid. For every pair of triangles, the line of intersection between the two planes must be found. If this line intersects either triangle then there are restrictions placed on the respective local coordinates.

The mathematics for these constraints proceeds in almost the same way as the previous section, except, in this case, the plane that is intersected with a triangle is simply the plane of the other triangle in the pair. This means the plane definition is not dependent on the position on the 'viewing' triangle, which means p and q may be set to zero in the constraint inequalities and that π_1 is already known. When the algebra is performed, there are only three coefficients in the resulting inequality in s and t.

Constraints of the same type are imposed by situations such as that shown in Fig 6.8. In this figure the triangles have been given new shapes to better illustrate why this type of constraint occurs. First, consider what parts of the light grey triangle are visible from point P_C on the dark grey

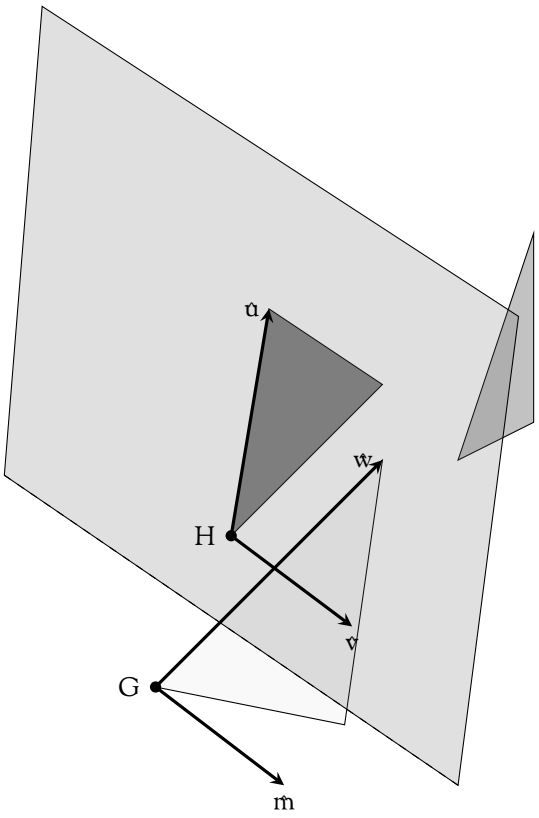


Figure 6.7: Visibility of the light grey triangle from the dark grey triangle blocked by the medium grey triangle.

triangle. The medium grey blocker would cause some of the tip of the light grey triangle near where $s = 1$ to be obscured. This type of obstruction was already handled. However if P_E was selected, and planes were constructed through the edges of the medium grey triangle, as discussed earlier, large portions of the light grey triangle would be "blocked", meaning that there are regions where a point in the region is enclosed by the planes that were constructed, but this is erroneous because no blocking can occur until the plane of the medium grey triangle is reached. The intersection line between the medium grey triangle and the plane of the light grey triangle is shown as a dashed line in Fig 6.8. On the near side of this line, no blocking may occur.

Type 3 constraints

The observation point must actually lie inside the dark grey triangle, and the "visible point" must actually lie inside the light grey triangle. Constraining p, q, s and t to $(0, 1]$ restricts points to the bounding rectangle of each triangle. Further constraints must be placed to limit the point to be inside the triangle rather than just the bounding square. These constraints depend on the direction and length of two shortest edges of each triangle. In Fig 6.9 an example of a dark grey triangle is shown in its local coordinate system. Assuming p and q are already constrained to $(0, 1]$ all that

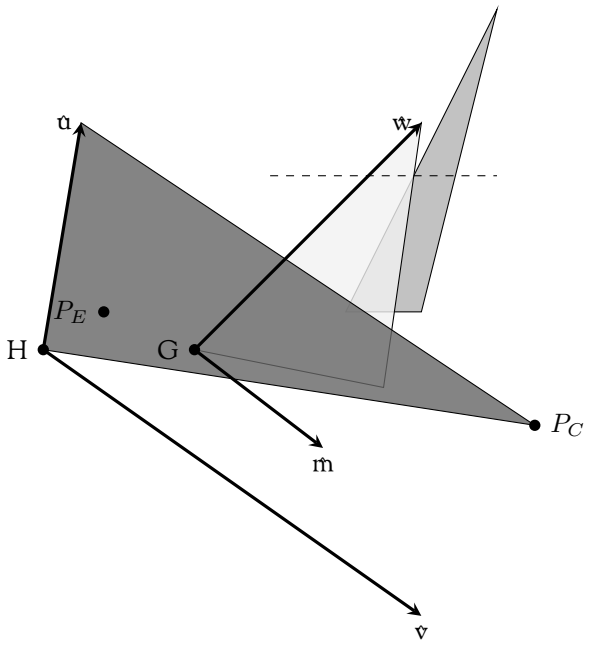


Figure 6.8: Visibility of the light grey triangle from the dark grey triangle blocked by the medium grey triangle.

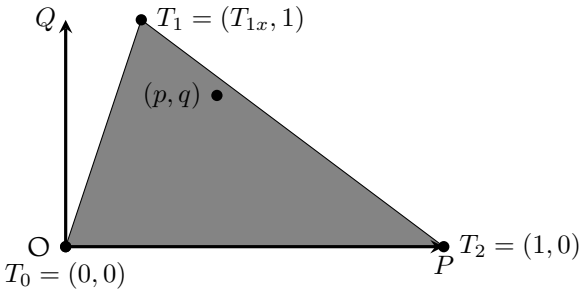


Figure 6.9: A triangle in it's local coordinate system

remains is to add constraints for the two shortest edges. The edge T_0T_1 gives the constraint

$$q < p \times \frac{1}{T_{1x}}, \tag{6.18}$$

or

$$q - \frac{p}{T_{1x}} < 0, \tag{6.19}$$

and the edge T_1T_2 gives the constraint

$$q < \frac{p}{1 - T_{1x}} - \frac{T_{1x}}{1 - T_{1x}} + 1, \tag{6.20}$$

or

$$q - \frac{p}{1 - T_{1x}} + \frac{T_{1x}}{1 - T_{1x}} - 1 < 0. \tag{6.21}$$

Exactly the same analysis applies to the target triangle (light grey in Fig 6.4) giving two additional constraints of the form of (6.19) and (6.20) on s and t.

6.5.6 Improvements Using Plücker Coordinates

The method, as outlined so far, is functional, and has the advantage that it is numerically stable due to the way the shadowing planes are constructed. A problem with the presented technique is its complexity. Type 1 constraints dominate both storage requirements and computational cost at run time. The inequalities formed for type 2 and type 3 constraints have relatively few coefficients (max 3). The type 3 coefficients can be calculated per triangle and stored $\mathcal{O}(N)$. The type 2 coefficients can be calculated per pair of triangles $\mathcal{O}(N^2)$. However, the type 1 constraints must be stored per blocker per pair per edge $\mathcal{O}(N^3)$, and each of these inequalities requires 60 double precision coefficients. This is inefficient in terms of both memory and processing time. In fact, this is also a problem for the 2D method in [6], in which a polygon must be stored per pair of triangles, with a number of edges proportional to the number of blockers.

To address this problem, the following steps are taken:

1. Relax the requirement to store all constraints as a series of coefficients for an inequality in p, q, s and t. Allow some intermediate steps to be evaluated in software functions.

2. Use Plücker coordinates [21] to encode each constraint.

Doing (1) is perfectly allowable in MIDACO - which accepts black box software routines as constraints - and means that only the Plücker coordinates per triangle edge need to be stored $\mathcal{O}(N)$, not per triangle edge per blocker per pair. Doing (2) means ray primitive intersection tests are extremely fast even with the small amount of extra calculation that must be performed due to coefficients not being precomputed fully per blocker pair.

The mesh is processed such that each edge of each triangle is stored using its Plüker coordinates. Given S and E, the start-point and end-point of a line segment respectively, the Plüker coordinates of the line containing both points is given as

$$L = \{E - S : E \times S\}. \tag{6.22}$$

Given two lines

$$L1 = \{U1 : V1\}, L2 = \{U2 : V2\}, \tag{6.23}$$

the quantity

$$w = U1 \cdot V2 + U2 \cdot V1. \tag{6.24}$$

The quantity w indicates how the two lines pass one another in 3D space, either intersecting, passing clockwise, or passing anticlockwise.

$$\left\{ \begin{array}{ll} \text{intersect,} & \text{for } w = 0 \\ \text{pass clockwise,} & \text{for } w < 0 \\ \text{pass anticlockwise,} & \text{for } w > 0 \end{array} \right\} \tag{6.25}$$

If a ray segment passes all the edges of a blocker in the same way, then it intersects the blocker. Using the new form of the blocking constraints, only the indices of the triangles between each ordered pair of triangles per pair of triangles need to be stored. This is still an $\mathcal{O}(N^3)$ storage requirement, but storing one integer per blocker per ordered pair is a clear improvement on 180 double precision floating point numbers. In reality, there are many pairs that are completely mutually invisible, and between any pair there are usually very few blockers, and often none at all. Furthermore if a primitive is between two triangles, but is not visible itself from the triangle from which the visibility is being established, it need not be stored as a blocker. Making savings like this means only the minimum required blocking set to determine visibility between any ordered pair is stored. This result is that far fewer than N^3 integers must actually be stored. In the case of the test environment used in this paper, with 3789 faces, the number of stored integers is less than 0.1 % of N^3, which is only about 0.2 GB of data when stored as alphanumeric characters in a text file. In contrast the 8 GB of RAM on the test machine was exhausted when a 3D model of the environment had greater than 200 primitives when encoding type 1 constraints the way initially presented.

The constraint inequalities are used twice, once when MIDACO is used to search for a feasible region to determine full/partial visibility, and once when validating paths between partially visible primitives. With constraints stored in Plücker form, there is an inefficiency because the software routine must convert to 3D coordinates from our local coordinate instances, before forming the Plücker coordinates of the line. This was not necessary in the type 1 format where the inequalities were directly in p, q, s and t. However, since the intra-mesh visibility is determined only once for an environment and stored, this feasibility search is generally not performance critical. Furthermore, using the type 1 constraints in their original form was only feasible on very small meshes due to memory constraints. Where the mesh is small, any performance enhancement is barely desirable anyway since the search is very fast regardless of method. When validating paths, the Plücker version is more advantageous because there is no need to convert 3D ray-triangle intersection points into the local coordinate system (although this must still be done for the type 2 and 3 constraints), and the Plücker type constraints consist of fewer multiples and adds than the constraint equations in type 1 form.

6.6 Benchmarking

6.6.1 Analysis of Intra-mesh Visibility

The visibility of the scene from a particular primitive - as determined by MIDACO's feasibility search under the formed constraints - is shown in Fig 6.11. Few primitives are visible when compared to the total number of primitives in the mesh. The empirical CDF of mesh visibility - as a percentage of the whole mesh - is shown in Fig 6.12 as a solid black line. For 90 % of primitives

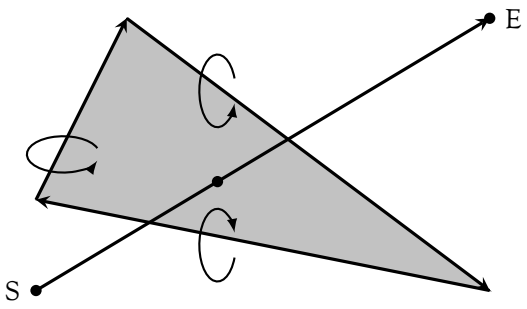


Figure 6.10: A line segment must pass all lines containing the edges of a triangle in the same way - either clockwise or anticlockwise - to intersect the triangle.

less than 10 % of the mesh is visible. For half of primitives, less than 5 % of the mesh is visible. For some special cases, e.g. the two triangles representing the ground plane and the vertical faces of the tallest buildings, very large proportions of the mesh are visible. This explains the gentle upper tail on the empirical CDF. The dashed trace, in Fig 6.12, represents the results obtained when only triangles that are behind, or have the same normal as the 'viewing' triangle are eliminated. This method neglects shadowing by other primitives, and many more triangles are counted as visible per triangle than for the more robust MIDACO method. This behind/backface-only method causes visibility to be overestimated. The result of this will be that there are many more nodes at each level of the path finding tree, resulting in more calculation at pruning stages of image tree generation, and more path-validation steps. These potential branches or paths would be known to be invalid *a priori* using the more advanced method using MIDACO. The behind/backface method, while simple and accurate (the same paths would be found eventually), creates unnecessary computation and therefore incurs a performance cost.

Unlike the conservative behind/backface method, some mesh visibility methods proposed in literature do not generate any false positives. However, due to a deficient analysis of visibility, these sometimes determine that primitives are mutually invisible when in fact they are mutually visible. The first is the simple circumcentre method. In this method, a ray segment is constructed between the centre of two primitives [9]. If the constructed ray segment intersects a blocker, then the primitives are treated as mutually invisible. This method is clearly liable to errors, especially when the primitive is large. A second method would be to avoid using the MIDACO library at all, and simply select variable instances (p, q, s, t) at random, using the same constraint equations to determine visibility/invisibility [25]. The benefit of this method is that the overheads associated with MIDACO are avoided, and the implementation is simpler. The cost would be that the advanced searching techniques implemented in the library are lost.

Fig 6.13 shows the mean percentage of the mesh visible from a primitive calculated using the methods discussed. It can be observed that the random and MIDACO methods converge to a mean mesh visibility of about 3.5 %. As the number of iterations on these methods increases to infinity,

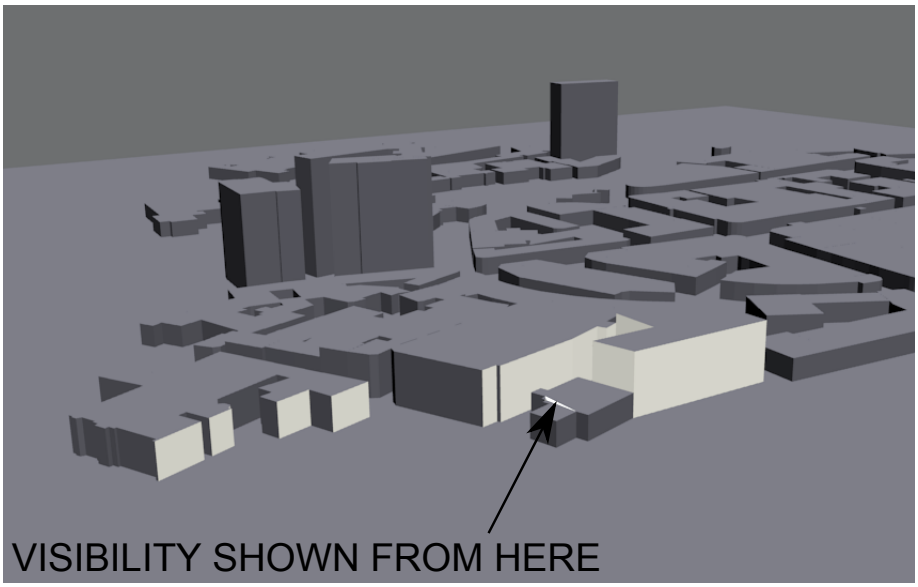


Figure 6.11: The primitives shown in medium grey are visible from the primitive indicated in light grey. Contains information from open street maps, which is made available at `https://www.openstreetmap.org` under the Open Database License (ODbL) `https://opendatacommons.org/licenses/odbl/`.

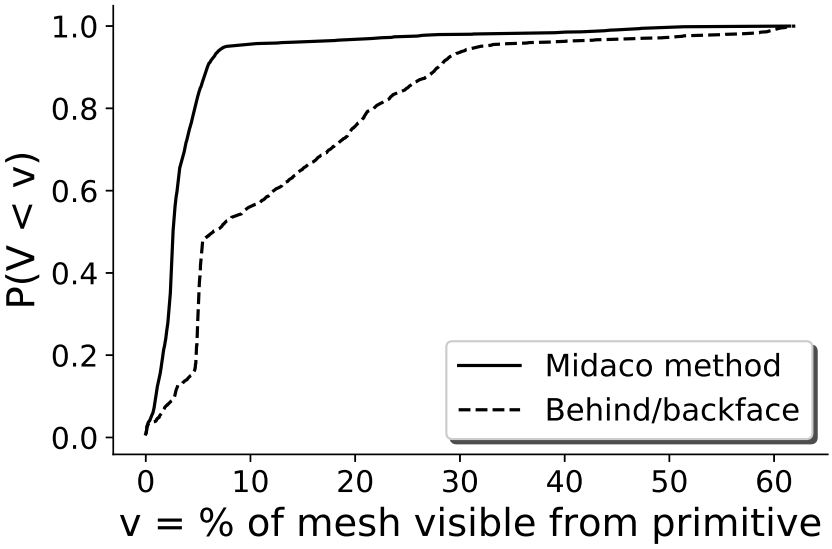


Figure 6.12: Empirical cumulative distribution function of mesh visibility from a primitive.

they will arbitrarily approach the exact solution. The number of iterations for these methods is indicated on the X axis. It is impossible for these methods to give a false positive. The number of false negatives decreases with increasing iterations. The plots start at 100 iterations, it can be observed that the MIDACO method appears to have converged by around 3000 iterations. Using

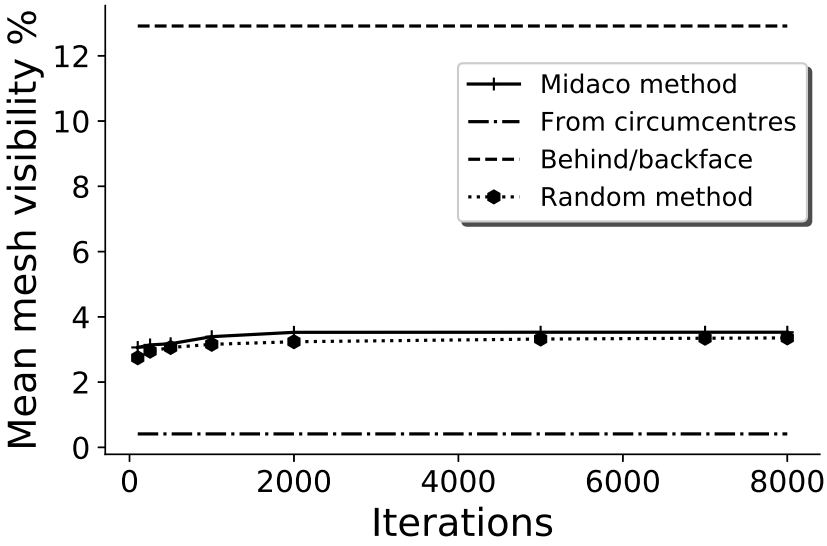


Figure 6.13: Mean percentage mesh-visibility per-primitive using the various intra-mesh visibility methods.

the backface/behind method gives many false positives. Over 12 % of the mesh is counted as visible from each primitive using the method. The circumcentre only method gives many false negatives, it is essentially the random method with only one iteration where the 'random' position on each triangle happened to be the circumcentre. As such, much less than 1 % of the mesh is visible from each primitive. The missing visible primitives are false negatives. False negatives will actually result in faster run times, but crucially give incorrect results!

Fig 6.14 depicts the same data for just the MIDACO and random methods to better illustrate their convergence. Per iteration, MIDACO converges much faster than the random method due to the advanced techniques used in the library. However, one iteration of the random method is about 4.86 times faster (mean, based on 7×10^6 iterations on the test computer) than an iteration of MIDACO's search. The time scaled version of the random method has also been shown in Fig 6.14. This means for the time scaled trace the number of iterations is 4.86 times the value on the X axis. This plot shows that, even when this time difference is accounted for, MIDACO converges more quickly to a stable solution at just over 2000 iterations, whereas the random method, which is able to do more than 8000 iterations in the same time, has not yet converged.

6.6.2 Performance Comparison

To analyse performance of the ray tracing tool, a 200 m by 200 m area of the test environment was installed with 52 transmitter locations. The open street maps data had flat terrain, although this is a general feature of the data set, it mirrors reality as the area is a very flat shopping precinct. Ran-

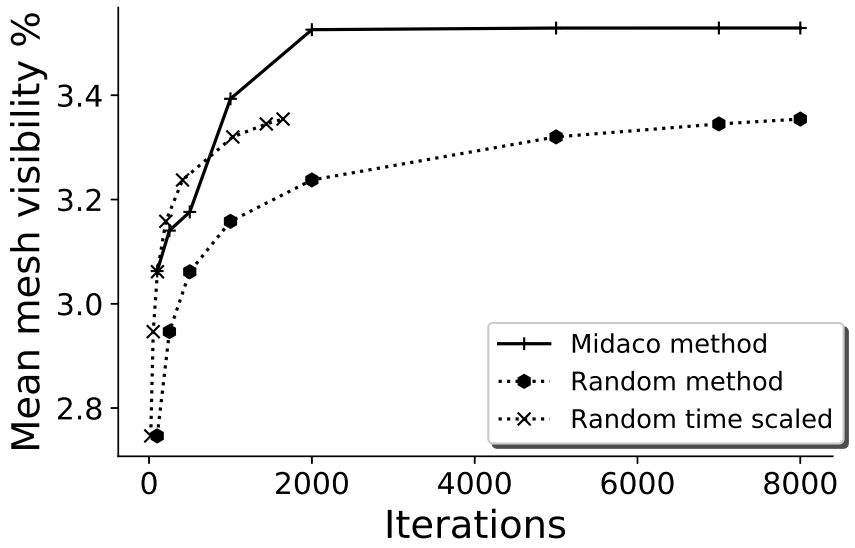


Figure 6.14: Mean percentage mesh-visibility per-primitive showing the MIDACO, random, and time scaled random methods.

dom user locations were also generated in this area, and then these were filtered so that only those outside buildings - i.e. in the street - remained. In total 1010 user locations were included in simulations. Path finding was conducted between each of the 52 base-station locations and each of the 1010 users giving 52520 individual channels in total. To benchmark path-finding using the visibility matrix found by the MIDACO method, it was compared it to that found by the behind/backface method. The method is close to the benchmark method used in [6]. Because the behind/backface methods visibility matrix is conservative, it gives identical actual results from path finding. The simulation was conducted finding reflections only. Various maximum reflection orders were stipulated for comparison. Fig 6.15 shows a particular BS UI pairing, with multiple found ray paths shown between them. Visibility of the mesh from the UI nodes was precomputed, but visibility from the BS nodes was included in the run times. Table 6.1 shows the results. It can be seen that

Table 6.1: Results

Interaction order	Links per Second (L/s) MIDACO	Time per Link (L/s) backface	Percentage improvement (%)
1	156	156	0
2	142	118	20
3	16.6	8.79	89
4	1.23	0.27	456

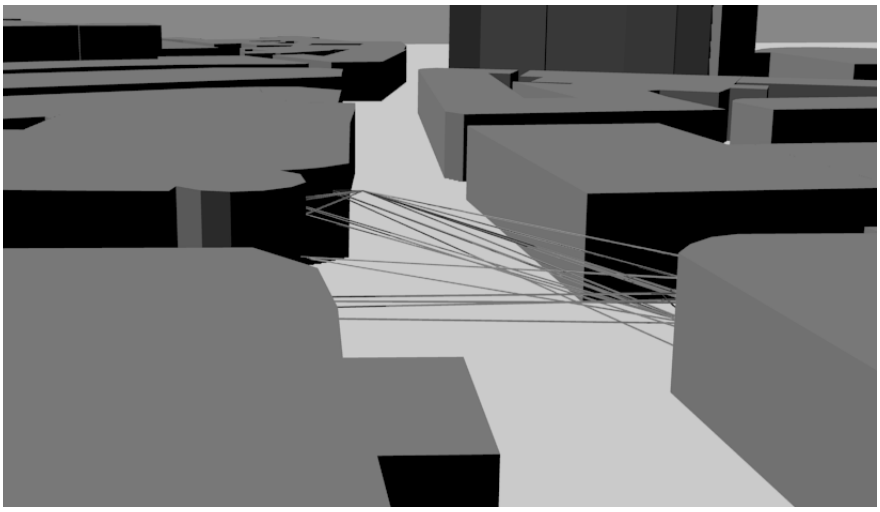


Figure 6.15: Ray paths for a BS-UI link. Contains information from open street maps, which is made available at `https://www.openstreetmap.org` under the Open Database License (ODbL) `https://opendatacommons.org/licenses/odbl/`.

for low maximum interaction orders the improvement is slight. This is because the simulation time is dominated by finding the initial visibility from the BS locations, not tree traversal. For a maximum interaction order of three, the optimised version is almost twice as fast as the basic version. When up to four interactions are allowed, the optimised version is over four times faster.

6.7 Further Analysis and Techniques

6.7.1 Analysis of Impact of False Negatives

Whilst the MIDACO method is much faster to get an output than the backface method, both methods do in fact find exactly the same ray paths. This could be checked by inspecting output files. Using the backface method means that false positive visibility relations are stored, this means unnecessary checking is performed, but no paths are missed because there are no false negatives. Many state of the art ray tracing tools use a method that generates false negative visibility relations. An interesting area of research is the impact on results if some false negatives are present.

False negatives increase execution speed, because fewer paths need to be checked, but can cause valid ray paths to be missed. As an illustrative example, Fig 6.16 shows some ray tracing results for the same link as Fig 6.15, except, in Fig 6.16, the visibility relations have been predicted using the circumcentre method, rather than the MIDACO method. Fewer ray paths are found between the BS location and the UI location. Here an analysis of the error in predicted power due to missed paths is provided.

It is worth noting here that the precision of the circumcentre method is dependent on the dis-

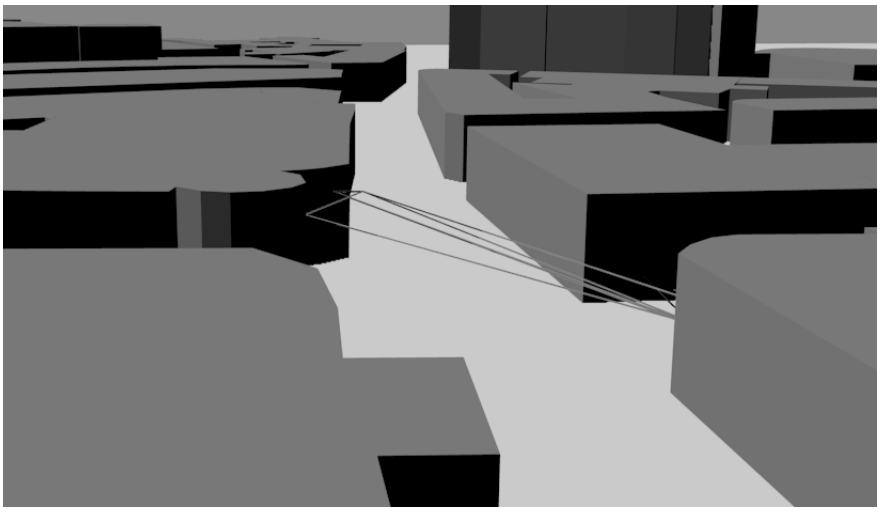


Figure 6.16: Ray paths for a BS-UI link. Contains information from open street maps, which is made available at `https://www.openstreetmap.org` under the Open Database License (ODbL) `https://opendatacommons.org/licenses/odbl/`.

cretisation of the environment. As faces are subdivided into more and more primitives, the intra-visibility relations would approach the perfect solution because the visibility to the centre of a very small primitive will be very representative of the visibility of the whole, very small, area. However, subdividing would greatly increase the number of primitives in the environment. Since the time taken for mesh preprocessing and the ray tracing operation itself is proportional to the number of primitives, this is not desirable. Here, the circumcentre method is used on a triangular mesh where the minimum number of primitives is used - i.e. a rectangular face is represented as two triangular primitives, and not more, regardless of its size.

For the simulations, which were introduced in the previous section, all the UI locations had an omnidirectional antenna radiation pattern applied, while the BS locations had a pattern for a patch antenna applied, with a directivity of 7 dBi. The main beam of the patch antenna pattern was angled 15 degrees towards the ground, and 30 degrees away from the closest wall. UI locations were 1 m from the ground while BS locations were 6 m from the ground.

Out of the 52520 individual links, 40808 had no paths found by the software when using either method. This was because, although every UI had links to several BS locations, most BS locations were far out of sight for a given UI location. This meant that after a maximum of three bounces no paths were found for that particular link. Of the links for which at least one ray path was found for at least one of the methods, 39.8 % of predictions were exactly the same using both methods. A further 17.7 % had at least one ray path found using one method, but none using the other. For these links the average predicted power for the link with at least one ray path was -106 dB. For the remaining 42 %, where both links were found to have some ray paths, but where one had more than the other, the average power difference (intensity) was 0.9 dB. Of links where there was no

LoS path, the average difference was 2.3 dB.

Whether these differences are important will be situation specific, but this result suggests that handling visibility correctly does have an impact on predicted power, and in certain instances will affect whether a ray path is present at all. These results quantify the benefits of using a robust method for intra-mesh visibility. Since less robust intra-mesh visibility methods are both easier to implement and cause faster run times, it is difficult to recommend one over the other without contextual information about the intended use of the extracted data.

6.7.2 Exploiting Proximity

So far in this Chapter, optimisation has focused on visibility relations between primitives in the environment. However, it was thought that further optimisations could be made when simulation MIMO networks by exploiting the fact that elements on a MIMO array are relatively close to one another within the wider environment.

There is little current research emphasis on classical MIMO in which communication takes place between two arrays with similar number of elements. Instead multi user MIMO schemes have been proposed. In these, a single large antenna array communicates with many single antenna terminals. It is assumed in this subsection that a single BS array will be simulated with a number of distributed UI locations that is greater than the number of array elements. Even though some MIMO systems are designed to have more array elements than users, this is a reasonable assumption because it allows statistically meaningful data to be pulled form a large data set retrieved from ray tracing data.

In current literature, each element on the array would be treated at an initial root node for an image search graph (see Fig 6.1). There would be multiple RX nodes in the graph, as shown in Fig 6.1, but only one TX node each time, from which the whole graph would be generated. There would be as many graphs as elements. What this corresponds to, in terms of actual process, is using each element array separately to generate image sources through the reflecting primitives. Because all the elements are quite close, many of the graphs would be quite similar. If, instead of reflecting a single element around to create image sources, it were possible to reflect around a whole array, multiple repetitions of the same section of graph would be avoided. Furthermore it would mean that the programme could benefit from some extra techniques which are presented here.

Routines to Optimise

In general, at each node on the path-search graph, two software routines are performed. The second (chronologically) of these routines performs frustum culling on the triangles that are candidates for further reflections. This routine uses an image source, and the primitive used to create the image source, to form the frustum. This happens for every node apart from those in the last layer in the graph; further reflections after the last layer are not allowed, deciding which triangles

would produce subsequently reflected paths is wasted effort. The other software routine is verification of ray paths. Verification happens for every array, for every UI location, for every valid combination of primitives making a route (i.e. every node in the graph). Because these two tasks are repeated so may times, they are the software routines in which most time is spent in the ray tracing application. This means that the optimisation of these routines is imperative to overall performance. Here, an optimisation of both tasks is presented using the relative proximity of elements in that array.

For Pruning

Classic Method The classic pruning operation is illustrated in two dimensions in Fig. 6.17. From facet H, the only invisible facet by pre-calculation is O which is blocked by M and N. By assessing visibility from the virtual transmitter after reflection through H, the only facets that are eligible for a further reflection are I, and J. By not creating further search paths for L, M, N and K, the total calculation has been reduced to one third of what would have otherwise been the case. At earlier levels in the graph, any pruning is particularly advantageous as large portions of the branching tree may be pruned. After multiple reflections the visibility region becomes particularly narrow, and this means more potentially visible primitives may be pruned, although for each pruned node the saving is smaller per pruning operation.

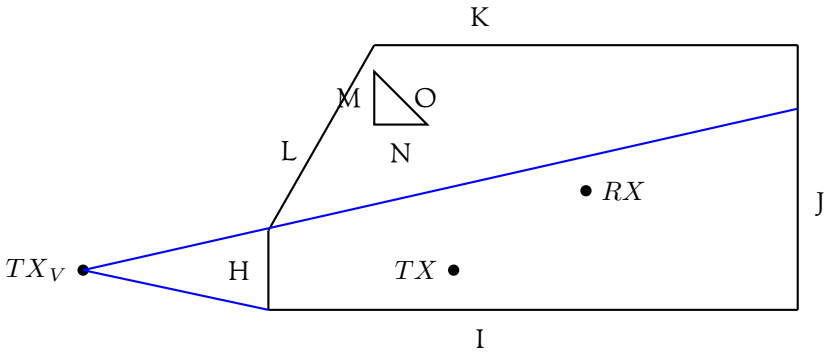


Figure 6.17: Pruning the image tree.

Array Method If an array, rather than individual point source, is considered, Fig. 6.17 is transformed the situation shown in Fig. 6.18. In 2D, it is easy to determine the visibility of remaining potentially visible facets. This process is illustrated with the purple and blue construction lines in Fig. 6.17. An analogous process may be performed in three dimensions. From each corner of an antenna array lines are projected through the corners of the primitive used to generate the frustum. For each edge of the primitive, the lines from the corner or the array that casts the widest view plane are used to define a clipping plane for that edge. By using the most conservative corner for each edge, it is possible to only reject triangles that are fully invisible from every location on the array.

Returning to the example, some leaves of the tree may be removed, as was the case for the single image source version. In this new case M, N and L are removed. However, K can now not be discounted because it is visible from some parts of the array. Initially, this seems like a disadvantage. For nodes on the array, such one positioned at the end of the array, from which the blue construction lines emanate, unnecessary path finding is created. Facet K would not have been visible if this node alone had been used to generate the image source. Clearly, to avoid this problem, some pruning must be performed for each location on the array individually. The advantage of the proposed technique is in noticing that many of the pruned primitives will be common to all elements on the array due to them being relatively localised. For L, M and N the pruning needs to happen only once for the whole array. If there were many elements on the array, performing the pruning once for the whole array would allow individual pruning to proceed much more quickly for the individual elements; there would be no repeat checking of facets that are invisible from all elements. In the example, when performing the path finding for the blue element, only the visibility of I, J and K and not L, M or N needs to be determined. If there were 10 elements on the array eliminating L, M and N would take 3 checks rather than 30. To summarise, by performing pruning for a whole array, the pruned primitives common to all elements can be dispensed with in a single shot. Some individual pruning is unavoidable, but by doing the common reduction, potentially large savings in computational effort may be made.

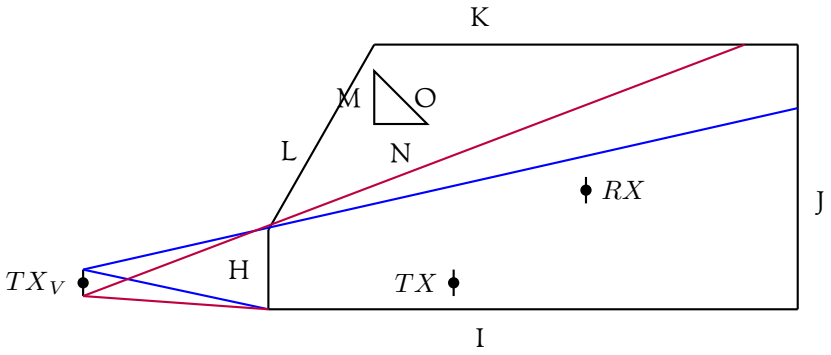


Figure 6.18: Pruning the image tree for a whole array.

If there are a total of t visible primitives that are candidates for further reflections after a reflection through a particular primitive, and the array has N elements, the proportion of t that is still visible after pruning for the 'whole' array image is x, such that for each individual element, after this initial cull, only xt facets need to be checked. The cost of a 'visibility from plane' check is C_p and the cost of a 'visibility from point' check is C_ϕ. For the proposed scheme to provide a benefit in terms of computational complexity it is required that

$$tNC_\phi > tC_p + xtNC_\phi, \tag{6.26}$$

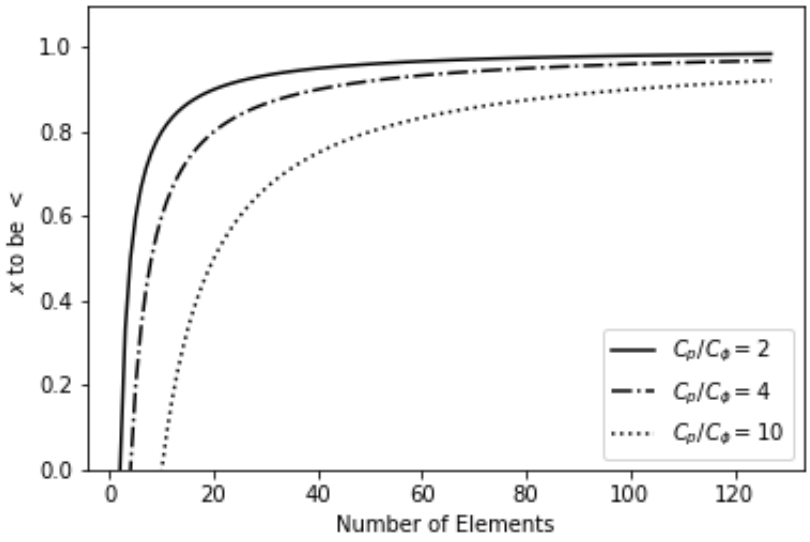


Figure 6.19: Viability of method by number of elements

or equivalently,

$$NC_\phi > C_p + xNC_\phi. \tag{6.27}$$

This means,

$$NC_\phi(1-x) > C_p, \tag{6.28}$$

so

$$-x > \frac{C_p}{NC_\phi} - 1, \tag{6.29}$$

or

$$x < 1 - \frac{C_p}{NC_\phi}. \tag{6.30}$$

If the ratio of time for a plane visibility to point visibility check, C_p/C_ϕ, is denoted as Z, then

$$x < 1 - \frac{Z}{N}. \tag{6.31}$$

This equation can be used to determine if, for various values of Z and N, a benefit may be achieved by the method, depending only on the proportion of possibly visible facets from the reflection facet still visible from a plane reflected through that facet, x. In Fig 6.19, the required value of x such that the proposed scheme offers a computational saving for various number of elements in an array is illustrated. The larger the number of elements, N, the easier it is for the scheme to provide a genuine saving. In 6.20 the number of elements is constant but the ratio C_p/C_ϕ is varied, showing that the longer a 'visible from plane' check is relative to a visibility check from a point, the more difficult it is to find an optimisation. However, in reality, C_p/C_ϕ will be about one. This is because the majority of both checks is pruning potentially visible triangles, rather than the initial frustum setup, which

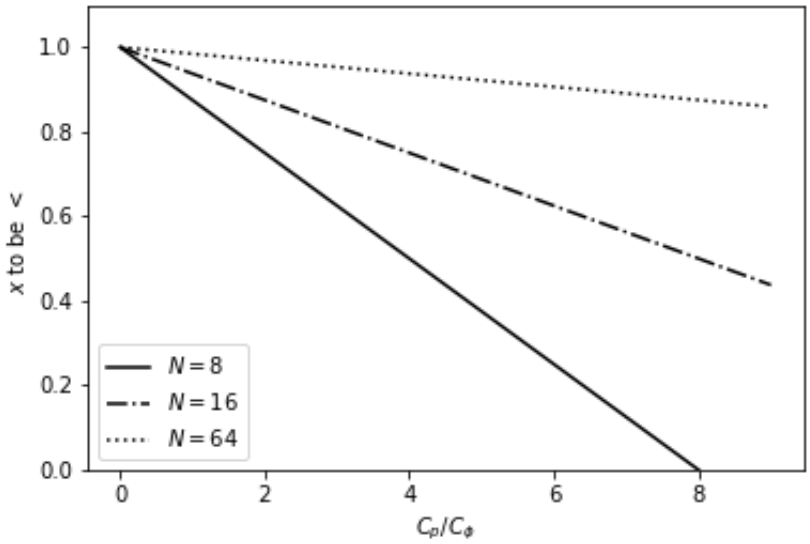


Figure 6.20: Viability of method by ratio of plane visibility cost to point visibility cost

is indeed longer for the 'from plane' check, but relatively insignificant when considered as part of the whole routine.

For Validation of Ray Paths

The proximity of array elements may also be exploited when validating ray paths. The process of validation, includes two main tests for every ray segment along the ray path. One is checking the ray path isn't intersected by another primitive, and therefore blocked. The other is that the intersection of a construction line between an image source and either an RX location or a previous intersect point in a multi-hop path, actually intersects the primitive used to generate the image. An excellent explanation with helpful illustrations is given in [2].

In Fig 6.21 a reflection of the array in a facet is shown. The facet is dashed to distinguish it. The paths to J are blocked but the paths to I are not. In fig 6.22 the paths to I intersect the facet used to generate the image - and hence are valid - whereas for J they do not. Checking for blockages may take considerable time, since there are potentially many blocking primitives with which intersection must be tested for. Checking the ray segment for an intersection with the primitive used to generate the image will be relatively quicker, because there is only one primitive to intersect with, and in many cases, this test will invalidate the path before any blocking needs to be investigated. Note that both the intersect point and whether a segment is blocked may be different for different elements on the array. However, it is possible to make statements about every element on the array based on calculating only a few points. Noting that all the elements are bounded by the elements in the four corners of the planar array, the corners may be used to establish if paths to the corners all

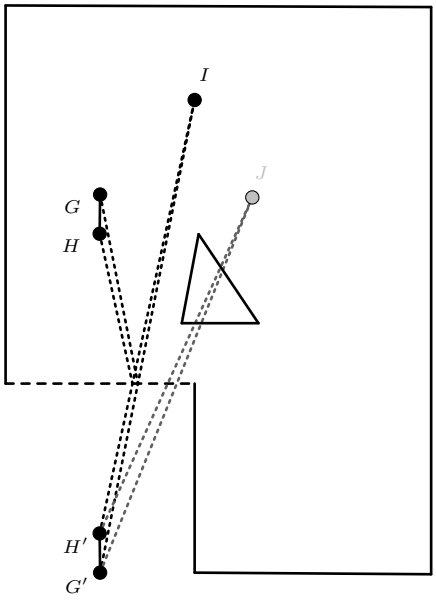


Figure 6.21: Being blocked.

fail to intersect a primitive used to generate a particular image in the same way. If all four corners' intersect-points lie outside the primitive by violating the same constraint, i.e. are off the same edge of the primitive, all the internal elements mat be counted as non-intersecting as well. In a multi bounce path this requires checking multiple primitives along the route in a loop. If it is found that some elements on the array do intersect every primitive along the route, then path checking proceeds as normal. Assuming this will be a minority of cases, then there is a possibility of massively optimising the routine by reducing the number of checks from as many elements as on the array to just four.

Performance

An implementation of the ray tracer using the proposed techniques was compared to the state of the art technique in which the graph is constructed separately for each array element - rather than making images of the whole array and keeping track of the elements that are still active after a certain number of reflections. A forty element MIMO array was set up in the test environment at rooftop height. It is reasonable to assume that simulation time scales linearly with the number of elements. Forty elements were used as this number is in the same order as common suggestions for real word array of 64 or 128 elements. A grid of 100 UI locations in a 10 by 10 m area was 1 m from the ground meaning there were 4000 individual links to simulate. Results are shown in Table 6.2.

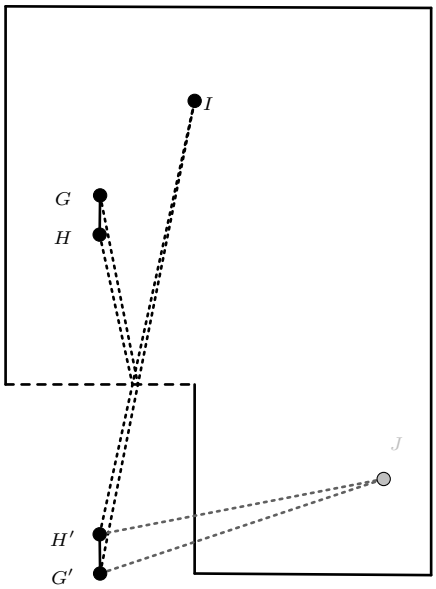


Figure 6.22: Not intersecting the reflecting primitive.

Interaction Order	1	2	3	4
State of Art (s)	250.1	305.8	2269	86670
Exploiting Proximity (s)	245.2	302.9	2041	108200

Table 6.2: Results comparing the state of the art to the new techniques.

The results presented in Table 6.2 are disappointing. It can be observed that only a marginal optimisation is made for orders of interaction up to three, and at four interactions use of the presented techniques actually slows down the execution when compared to the state of the art. As the number of allowed interactions increases, fewer array elements will be still active after a given series of reflections. This means that the optimisation techniques - which rely on their being may array elements - fail to produce significant savings, while the overheads associated with them consume valuable time. The use of profiling reveals that, because there is an exponentially growing number of nodes at each layer, graph construction and internal (i.e. not last layer) nodes actually consume relatively little of the total execution time. This means most of the time is spent in the last layer of nodes performing verification. This causes a great 'evening up' between the proposed method and the state of the art. Both methods will produce the same last layer paths needing verification, and since there are so many of these, any optimisations up to that point are insignificant in comparison to the calculation that must be performed for these last layer links. The method by which a whole set of ray paths can be rejected using only four points should still provide an optimisation, and indeed, of the two techniques presented, it is the one responsible for the optimisation seen on interaction orders 1 - 3. However, in some cases, it will be unable to reject a path, in which case it is wasted calculation. In other cases, only a few elements will still be active, if there are too few, no optimisation may be made and again computation can be wasted. Unfortunately the conclusion of the short study is that the proposed proximity exploiting techniques fail to provide a meaningful optimisation.

6.7.3 Validity of the Plane Wave Assumption

The plane wave assumption is commonly used to extrapolate amplitude and phase values across an array from a single shot of ray tracing to its central element. Using knowledge of the angles of arrival of the various ray paths, and of the orientation of the array, simple trigonometry may be used to calculate the difference in path lengths to different elements on the array. This information is used to adjust path loss and phase predictions appropriately.

To explore the validity of this method, a simple simulation was conducted. A horizontal line segment that was 2 m long and at roof height was sampled every centimetre. These points were counted as element locations on a 2m long linear array. The array transmitted to a point at street level, with a horizontal separation between the point and the linear array of about 50 m. Both the linear array and the point where arranged in a model of a street in Broadmead, Bristol. The arrangement of the array and the point was such that a LoS path between them was always present. Simulations were conducted at 3.5 GHz. The ray paths between an element on the array and the point at street level are shown in Fig 6.23. The array was oriented in the Y direction, which corres-

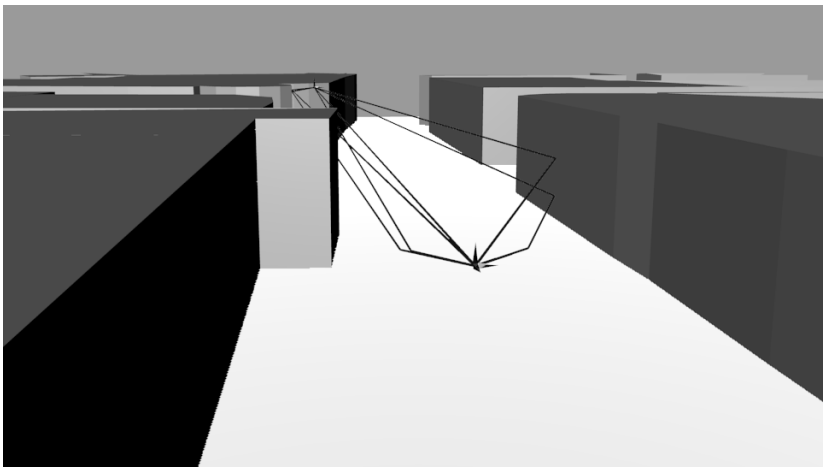


Figure 6.23: Ray paths between an array element and a point a street level. Contains information from open street maps, which is made available at `https://www.openstreetmap.org` under the Open Database License (ODbL) `https://opendatacommons.org/licenses/odbl/`.

ponds to travelling from left to right across the page when referring to Fig 6.23.

Brute force ray tracing was performed between every array element and the point at street level separately. All elements had an isotropic antenna radiation pattern applied. The results from the brute force ray tracing were compared to those in which the ray tracing results from the first element were used to extrapolate values for subsequent elements using the plane wave assumption. Results are shown in Figs 6.24 and 6.25, with the solid black line representing the brute force results and the dashed grey line representing the extrapolated results. It can be observed that as the element of interest moves further from the initial element, the extrapolated results deviate further and further from the more accurate brute force results. Inspection of the ray tracing results files shows one key contributor is the presence of extra ray paths that appear about 60 cm along the array from the initial element. It can be observed in Fig 6.24 that the fading behaviour changes subtly after this distance, with more frequent shallow fades superimposed on the previous fading pattern.

Missing paths is a problem that is intrinsic to the plane wave assumption extrapolation method. A trade off must be made between execution speed and accuracy since the extrapolation method will be faster than the brute force to a degree linearly proportional to the number of elements in the array. It is difficult to generalise the results presented because they are highly environment specific. However, what is clear is that, over a distance similar to plausible dimensions for a MIMO array at 3.5 GHz, there is significant variation in the multipath structure of the channel, and this means, even for a line of sight scenario, some errors are introduced. It is probable that in a NLoS scenario, with a less dominant main propagation path, the effect would be more noticeable. In situations where an array is half shadowed to a LoS component, again this type of error would be severe. Results for only 3.5 GHz are presented because this is the frequency at which new MIMO systems are expected to operate.

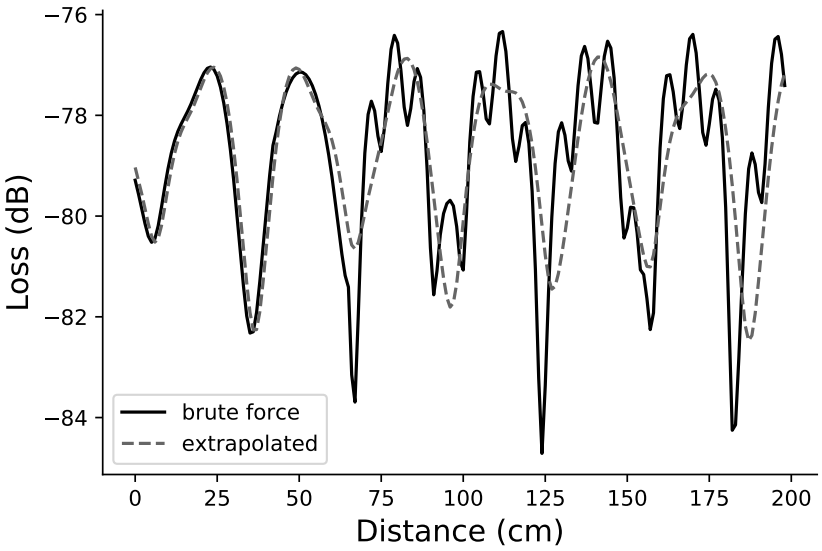


Figure 6.24: Loss as a function of distance along a short path representing an array. The plots show received power level at each point along the line from a single transmitter perpendicularly down the road from the line representing the array.

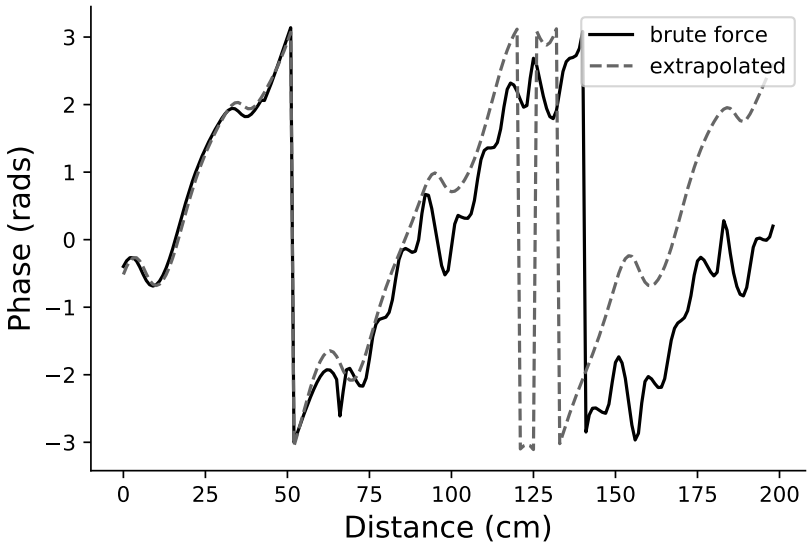


Figure 6.25: Phase.

Bibliography

[1] Zahid Aslam, Yoann Corre, Emil Björnson, and Erik G. Larsson. Performance of a dense urban massive mimo network from a simulated ray-based channel. *EURASIP Journal on Wireless Communications and Networking*, 2019, 12 2019.

[2] Henry L. Bertoni. *Radio Propagation for Modern Wireless Systems*. Prentice Hall Professional Technical Reference, 1999.

[3] M. Boban, J. Barros, and O. K. Tonguz. Geometry-based vehicle-to-vehicle channel modeling for large-scale simulation. *IEEE Transactions on Vehicular Technology*, 63(9):4146–4164, Nov 2014.

[4] V. Degli-Esposti D. Bilibashi, E. M. Vitucci. Dynamic ray tracing: Introduction and concept. COST IRACON, *7* 2019.

[5] V. Degli-Esposti, F. Fuschini, E. M. Vitucci, M. Barbiroli, M. Zoli, L. Tian, X. Yin, D. A. Dupleich, R. Müller, C. Schneider, and R. S. Thomä. Ray-tracing-based mm-wave beamforming assessment. *IEEE Access*, 2:1314–1325, 2014.

[6] N. Dreyer and T. Kürner. An analytical raytracer for efficient d2d path loss predictions. In *2019 13th European Conference on Antennas and Propagation (EuCAP)*, pages 1–5, March 2019.

[7] Christer Ericson. *Real-Time Collision Detection*. CRC Press, Inc., Boca Raton, FL, USA, 2004.

[8] S. J. Flores, L. F. Mayorgas, and F. A. Jimenez. Reception algorithms for ray launching modeling of indoor propagation. In *Proceedings RAWCON 98. 1998 IEEE Radio and Wireless Conference (Cat. No.98EX194)*, pages 261–264, Aug 1998.

[9] Franco Fuschini, Enrico M. Vitucci, Marina Barbiroli, Gabriele Falciasecca, and Vittorio Degli-Esposti. Ray tracing propagation modeling for future small-cell and indoor applications: A review of current techniques. *Radio Science*, 50(6):469–485.

[10] Franco Fuschini, Enrico M. Vitucci, Marina Barbiroli, Gabriele Falciasecca, and Vittorio Degli-Esposti. Ray tracing propagation modeling for future small-cell and indoor applications: A review of current techniques. *Radio Science*, 50(6):469–485, 2015.

[11] S. Hussain and C. Brennan. An image visibility based pre-processing method for fast ray tracing in urban environments. In *2016 10th European Conference on Antennas and Propagation (EuCAP)*, pages 1–5, April 2016.

[12] S. Hussain and C. Brennan. Efficient preprocessed ray tracing for 5g mobile transmitter scenarios in urban microcellular environments. *IEEE Transactions on Antennas and Propagation*, 67(5):3323–3333, May 2019.

[13] J. Järveläinen and K. Haneda. Sixty gigahertz indoor radio wave propagation prediction method based on full scattering model. *Radio Science*, 49(4):293–305.

[14] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta. Massive mimo for next generation wireless systems. *IEEE Communications Magazine*, 52(2):186–195, February 2014.

[15] J. S. Lu, E. M. Vitucci, V. Degli-Esposti, F. Fuschini, M. Barbiroli, J. A. Blaha, and H. L. Bertoni. A discrete environment-driven gpu-based ray launching algorithm. *IEEE Transactions on Antennas and Propagation*, 67(2):1180–1192, Feb 2019.

[16] Y. Miao, Q. Gueuning, M. Gan, and C. Oestges. Adding diffuse scattering correlation to effective roughness models in ray tracing. In *2017 11th European Conference on Antennas and Propagation (EUCAP)*, pages 828–830, March 2017.

[17] Y. Miao, Q. Gueuning, and C. Oestges. Modeling the phase correlation of effective diffuse scattering from surfaces for radio propagation prediction with antennas at refined separation. *IEEE Transactions on Antennas and Propagation*, 66(3):1427–1435, March 2018.

[18] Tomas Möller and Ben Trumbore. Fast, minimum storage ray-triangle intersection. *Journal of Graphics Tools*, 2(1):21–28, 1997.

[19] NEBENS. Mimobit. `http://www.nebens.com`, 2018.

[20] J. Nuckelt, M. Schack, and T. Kürner. Geometry-based path interpolation for rapid ray-optical modeling of vehicular channels. In *2015 9th European Conference on Antennas and Propagation (EuCAP)*, pages 1–5, April 2015.

[21] Helmut Pottmann, Martin Peternell, and Bahram Ravani. An introduction to line geometry with applications. *Computer-Aided Design*, 31(1):3 – 16, 1999.

[22] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez. Millimeter wave mobile communications for 5g cellular: It will work! *IEEE Access*, 1:335–349, 2013.

[23] Martin Schlüter, Matthias Gerdts, and Jan-J. Rückmann. A numerical study of midaco on 100 minlp benchmarks. *Optimization*, 61(7):873–900, 2012.

[24] Jonathan R Shewchuck. `https://www.ics.uci.edu/~eppstein/junkyard/circumcenter.html`.

[25] Eustace Tameh. *The Development and Evaluation of a Deterministic Mixed Cell Propagation Model based on Radar Cross-Section Theory*. PhD thesis, University of Bristol, 1998.

[26] F. Wiffen, L. Sayer, M. Z. Bocus, A. Doufexi, and A. Nix. Comparison of otfs and ofdm in ray launched sub-6 ghz and mmwave line-of-sight mobility channels. In *2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pages 73–79, Sep. 2018.

[27] Z. Yun and M. F. Iskander. Ray tracing for radio propagation modeling: Principles and applications. *IEEE Access*, 3:1089–1100, 2015.

Chapter 7

Mobility

Some results in this chapter were obtained in collaboration with colleagues in the communications systems and networks group at the University of Bristol. Specifically, section 7.2.8 contains results from bit level simulations that were performed by F. Wiffen as part of the co-authored work [25] and some of the introductory text to the orthogonal time frequency space waveform is also included from the co-authored work. The paper [25] is almost completely reproduced including figures and tables.

7.1 Dynamic Ray Models

Ray modelling is normally performed on static scenes and indeed many practical environments are largely stationary. This is not always the case however. In fact, vehicular networks - which are one of the desired use cases for 5G networks - will become ubiquitous with the uptake of autonomous cars and improved access to broadband on trains. These are good examples of non-static propagation environments. There are two main reasons why it is important to treat mobility properly in a ray based propagation model. The first is that the geometry of the scene changes with time, affecting the channel, and the second is that mobile objects and UEs cause Doppler shifts which can have consequences, for example, if orthogonal frequency division multiplexing (OFDM) inter-carrier spacing is small relative to the Doppler shifts in the channel, this can cause throughput degradation [10].

7.1.1 Updating the Scene

A scene changes due to the motion of objects in it causing the channel to evolve with time. If the same snapshot of a static scene is simulated several times, identical results will be produced at each "time step" and the realistic evolution of the channel would not be captured. The updating of the scene at each time step could be performed manually, and simulations be performed on several meticulously prepared scenes, but this is tedious. The coherence time and distance can be low in wireless communications systems. If a realistic channel progression is required, the updating of the scene must be in increments that are less than the coherence time and distance. The process may be automated such that only the starting positions and the velocities of transceivers and objects need to be specified.

This was implemented in software by using a property of the way the triangular mesh representing a scene was stored. The mesh was stored as a list of vertices followed by a list of "facets" comprised of triplets of indices of the vertices already specified and tagged with material types. Velocities were added as a property of a particular material. By assigning separate materials to all objects with different material properties or velocities, the programme was able to identify which vertices need to be update at each time step. Because the velocity was known for each material type, how much to update each vertex by is based on a predetermined time step. The scene was then updated in the software using the new vertex positions. By storing moving parts of the scene in a separate data structure, only those that have been updated need to be committed to the acceleration structures used in software each time the scene is updated.

7.1.2 Doppler shifts

The second reason for paying proper attention to the motion of objects in the scene is for the investigation of Doppler shifts. Whilst the Doppler shift on a ray path is relatively straight forward to calculate if the velocities of the objects that are hit along the path are known, Doppler is difficult to derive from measurements, and is often treated overly simplistically in statistical propagation models. A method for calculating Doppler shifts on ray paths is presented here.

Considering Fig 7.1, a propagation path between two moving user equipments (UEs) is shown. The ray path is made of several segments, each a straight line between a source point and a destination point. The object at both the source and the destination point (either a UE or primitive) has an arbitrarily oriented velocity in space. In general the source velocity is independent of the destination velocity, and both are independent of the direction of the ray segment between them. The Doppler shift over a segment can be calculated using (7.1).

$$f_{n+1} = \left(\frac{c + \mathbf{v}_d \cdot \mathbf{s}_n}{c + \mathbf{v}_s \cdot \mathbf{s}_n} \right) f_n. \tag{7.1}$$

In (7.1), f_n is the frequency at the start of the ray segment, $\mathbf{v}_d$ and $\mathbf{v}_s$ are destination and source velocity vectors respectively, $\mathbf{s}_n$ is a unit direction vector describing the direction of propagation along

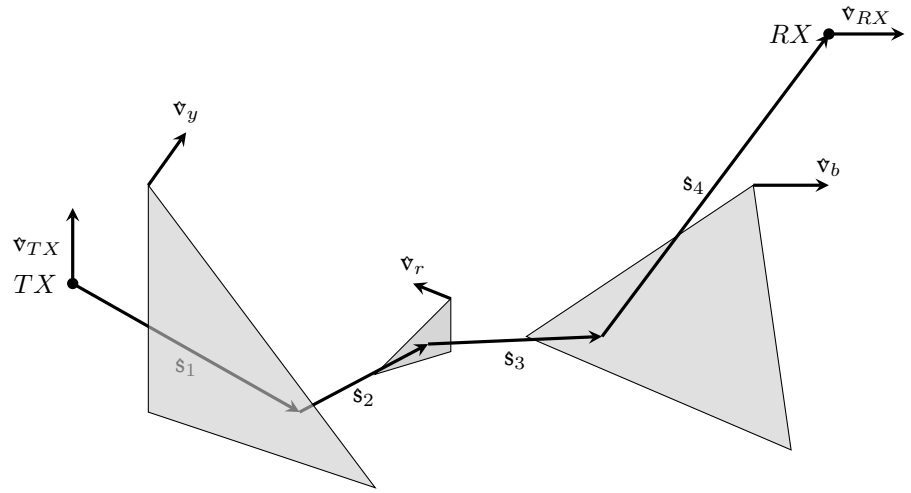


Figure 7.1: A ray path from a moving TX, bouncing off three moving primitives arriving at a moving RX.

the n^{th} ray segment, c is the speed of the wave in the medium and f_{n+1} is the apparent frequency at the destination due to the Doppler shift caused by the relative motion, along the direction of propagation of the ray, of the source and destination objects. With f_0 set equal to the carrier frequency, (7.1) can be used in conjunction with (7.2) to calculate ξ_i, the Doppler shift on the i^{th} ray path, where the i^{th} path had m ray segments.

$$\xi_i = f_m - f_0. \tag{7.2}$$

7.2 Comparison Of OTFS and OFDM in Ray Launched Mobility Channels

Orthogonal Time Frequency Space (OTFS) [7] is a recently proposed modulation scheme for doubly-dispersive channels in which symbol multiplexing and processing is performed in the Doppler-delay domain, rather than conventional time-frequency domain. In this section, the performance of OTFS is compared to orthogonal frequency division multiplexing OFDM for line-of-sight mobility automotive channels.

Ray launching is used to simulate the channel for two different dynamic 3D vehicle to infrastructure transmission environments, using a Kirchhoff model for diffuse scattering from rough surfaces. Bit level simulations for transmission from a transmitter moving at speeds of 13 m/s and 31 m/s were then carried out, for both OFDM and OTFS. It was found that that with short length block codes OTFS outperforms OFDM in all simulated scenarios, reducing the block error rate by more than 50% on average.

7.2.1 Background

With 5G communication services aiming to provide ubiquitous network access, there is a growing need for communication systems capable of providing good quality of service wireless access in high mobility environments, i.e. to and from moving vehicles. These mobile environments are characterised by having a wireless channel that is doubly dispersive - spreading signals across both time, due to the delay of the propagation paths, and frequency, due to Doppler shifts caused by motion in the channel. This is problematic for conventional waveforms such as OFDM, which suffer from heavy fading in a rapidly varying channel.

Recently, the OTFS modulation scheme was proposed as a solution to these issues [6]. The main innovation of the OTFS waveform is that, unlike conventional waveforms, which operate in the time-frequency domain, in OTFS symbol multiplexing and detection is performed across a grid in the 'Doppler-delay' domain. Viewed in this domain, the channel response is decomposed into a slowly varying response that closely reflects the physical geometry of the channel. A Doppler-delay impulse transmitted through the channel appears after demodulation as a series of impulses translated according to the delays and Doppler shifts of the constituent propagation paths. Providing the resolution is sufficient to separate the received signal components, fading is eliminated and the full diversity of the channel is captured with each transmitted symbol experiencing the same channel gain (assuming ideal pulse-shaping). In addition, OTFS can be implemented as a pre- and post-processing stage to standard multicarrier modulation schemes, enabling it to co-exist with other 5G technologies.

Here, a performance comparison of OTFS and standard OFDM is provided - at both sub-6 GHz and mmWave - using ray launched line-of-sight channels captured for two automotive mobility scenarios namely a built up urban environment and a motorway environment.

7.2.2 Ray Launching Simulations

Ray launching was used to simulate the propagation channel between a mobile station and a base-station at 3.5 GHz and 28 GHz, designated frequency bands for future 5G systems. Rays were fired "equally" from a source in three dimensions. 10,000 rays were used, corresponding to an angular resolution of 0.035 rads. See section 4.3. This resolution was used as a trade off between directional accuracy and simulation time.

7.2.3 Power of Individual Paths

In general, each ray path between TX and RX consists of one or more ray segments. The received power of the i^{th} ray path, $|h_i|^2$, is calculated by multiplying linear power gains and losses along that path. This includes antenna gains at the TX/RX $G_{TXi/RXi}$, losses due to reflection/scattering for the j^{th} interaction, $K_{[i,j]}$, and losses due to propagation in free space, P_i. Apart from P_i, these are

2×2 matrices specifying co- and cross-polar coefficients for vertical and horizontal polarisation. For the simulations presented here, the transmitted power was vertically polarised and only the vertical channel was observed at the RX. Rotation matrices, $\mathrm{R}[i, j]$, were used to resolve a polarisation vector into a local vertical and horizontal polarisation before interaction with a primitive, as in [15]. The total loss for a given ray path is given by (7.3) where $\mathrm{b} = \begin{bmatrix} 1 & 0 \end{bmatrix}^T$.

$$|h_i|^2 = \mathrm{b}^T \mathrm{G}_{TXi} \mathrm{R}_{[i,0]} \mathrm{K}_{[i,0]} .. \mathrm{K}_{[i,m]} \mathrm{R}_{[i,m+1]} \mathrm{G}_{RXi} \mathrm{b} P_i \tag{7.3}$$

where the propagation path loss was calculated according to (7.4) with d as the distance for the ray path in metres and λ is the wavelength in metres.

$$P_i = \left(\frac{\lambda}{4\pi d} \right)^2 \tag{7.4}$$

In the case that the j^{th} interaction is a specular reflection, $\mathrm{K}_{[i,j]}$ corresponds to the Fresnel reflection matrix, reduced by a factor to account for the roughness of the surface [4]. This reduction factor is shown in (7.5) below.

$$g(\sigma) = \exp\left[- \left(\frac{4\pi\sigma \cos(\theta)}{\lambda} \right)^2 \right] \tag{7.5}$$

where λ is the wavelength in metres, θ is the angle of incidence, and σ is the standard deviation of the height of the surface in metres.

$\mathrm{R}_{[i,x]}$ is a rotation matrix to decompose the incoming polarisation vector into local transverse electric and transverse magnetic mode for the x^{th} interaction. A formulation for this is given as,

$$\mathrm{R}(\varphi) = \begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix} \tag{7.6}$$

[16], where φ is the angle between the local transverse electric mode for the scattering element, and the transverse electric mode for the ray, around the direction of the ray.

The phase of each h_i is the cumulative phase shift from the reflections along that path,

$$\angle h_i = \phi_{TXi} + \phi_{i,0} + \ldots + \phi_{i,m} + \phi_{RXi}. \tag{7.7}$$

7.2.4 Scattering

At 28 GHz diffuse scattering is an important propagation mechanism. To simulate this mechanism, a version of the Kirchhoff model [8]-[23], which was verified in [22], was used in the ray launching tool. The ray model simulated up to five reflections and one diffuse scatter, considering diffuse scatters only for either the last interaction before arriving at the RX, or the first interaction after leaving the TX. In the case $\mathrm{K}_{[i,j]}$ was the matrix for a diffuse scatter it was formulated as shown in [15]-[4],[13]. The area of a scattering tile is calculated based on the distance from the source and angle between rays.

Material	ϵ	σ (m)	Ω (m)
Metal	$9 + 3000j$	0	0.1
Concrete	$3 + 3j$	4.1×10^{-4}	1.3×10^{-3}

Table 7.1: Properties of the materials considered in simulations.

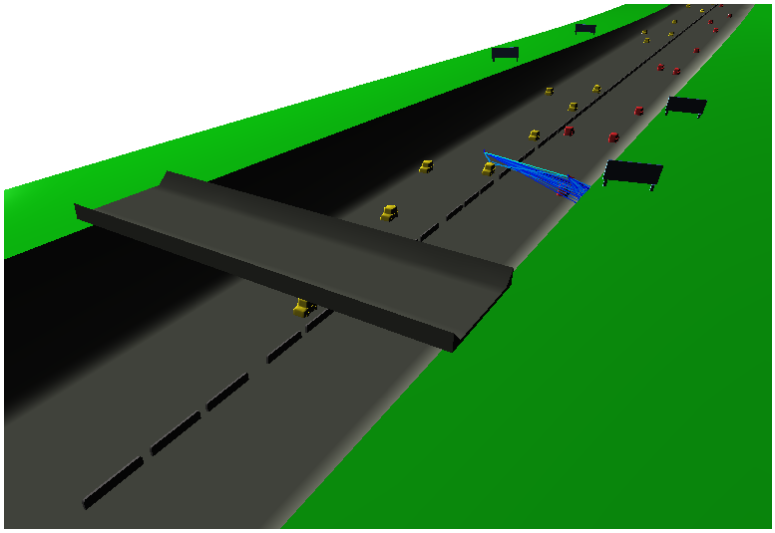


Figure 7.2: Motorway scenario.

7.2.5 Dynamic Scenarios Simulated

Two scenarios were simulated. The first was a section of motorway as shown in Fig. 7.2, with bridges, billboards and crash barriers included. The other scenario was a city, based on a small section of London, England, as shown in Fig. 7.3. The cars and billboards were made of metal and the road surface, bridges, buildings and crash barriers were concrete. Table 7.1 shows the parameters used for these materials including relative permittivity, ϵ, and correlation length, Ω, which is defined in [8][20]. For metallic surfaces diffuse scattering was not simulated due to their smoothness.

In both cases a straight route of 40 m was traversed in 2 m steps by the TX, with the RX being a stationary base station. In the motorway the base-station antenna was at a height of 9 m and was 5 m across the road from the TX, which was at a 1.1 m height. The TX started 30 m away from the BS, moving past the base-station at a velocity of 112 km/h (31 m/s). In the town the base-station was at a 15 m height, a height similar to the rooftops. The TX was at 1 m height and started at a position 15 m from the base-station moving along the street away from it at a velocity of 48 km/h (13 m/s). The locations of the cars were updated based on their velocity at each time-step in order to properly represent the dynamic scene. At all simulated locations, a clear line-of-sight path between TX and RX existed.

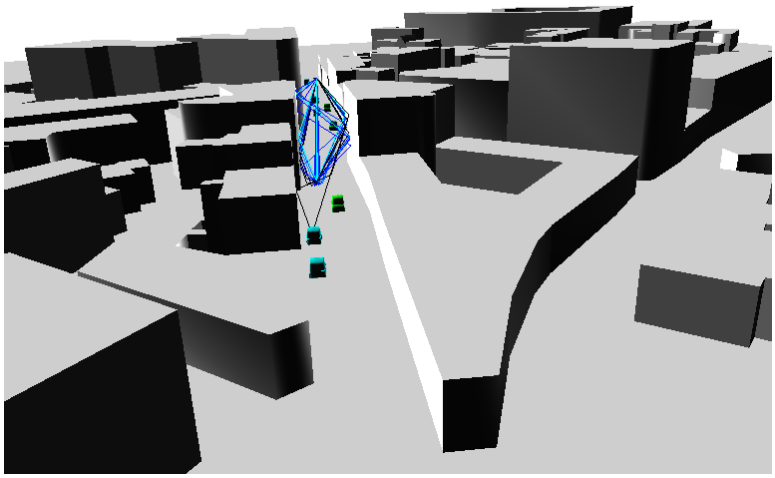


Figure 7.3: Town scenario, contains information from open street maps, which is made available at `https://www.openstreetmap.org` under the Open Database License (ODbL) `https://opendatacommons.org/licenses/odbl/`.

7.2.6 Doppler Shifts

In existing work it is assumed that many propagation paths will exist, with uniformly distributed angle of arrival and thus random Doppler shifts. This is unrealistic for real line-of-sight channels. In Fig. 7.4 the Doppler power spectrum for the channel at 28 GHz on the motorway is shown for each position on the route. This should be contrasted to the classic "bathtub" Doppler power spectrum of Jakes's model.

7.2.7 Antenna Patterns

For the 3.5 GHz simulations an omnidirectional pattern was used. However, directional antennas are generally used in millimetre wavelength communication systems to mitigate interference, overcome propagation losses and blockage caused by moving objects [18]. The antennas have the effect of spatially filtering the channel. A simple directional "pattern" with a beam-width of 15° was applied according to (7.8), to point directly "along" the strongest propagation path for channels captured at 28 GHz. Although the ray tracer had the capability to use more realistic patterns, it was decided that this pattern was an acceptable approximation to a generic directional antenna, or antenna array to be used in mmWave systems.

$$G_i = \begin{cases} 31.6, & \theta_i \leq 7.5^\circ \\ 0.87, & \text{otherwise.} \end{cases} \tag{7.8}$$

In (7.8), G_i is the linear antenna directivity for the i^{th} ray and θ_i is the angle between the i^{th} ray

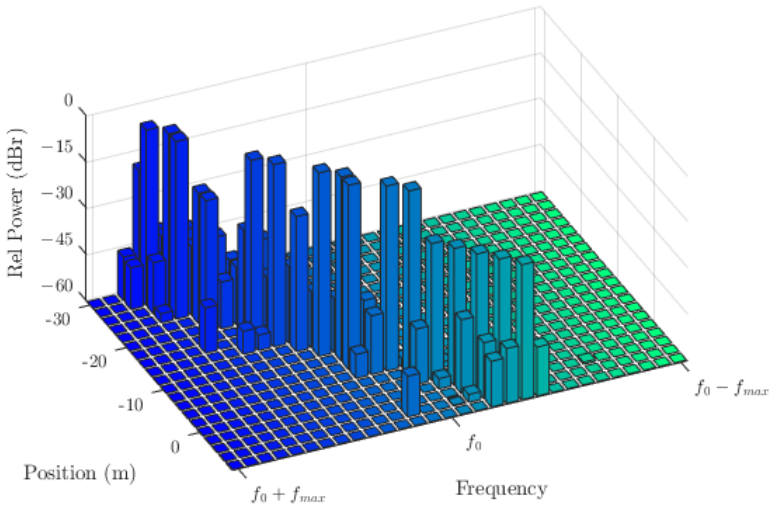


Figure 7.4: Doppler power spectrum along the route at 28 GHz.

f_0	Δf	M	N
28 GHz	60 kHz	1667	7
3.5 GHz	30 kHz	667	14

Table 7.2: Simulation configurations.

and the strongest omnidirectional ray. These linear directivities correspond to 15 dB and -0.6 dB. The gains prescribed by (7.8) were applied separately at both the TX and the RX for each ray path. This pattern is shown in Fig. 7.5.

Whilst the moving TX antenna was floating in free space in simulations, it was supposed to represent a roof mounted antenna on a car so any path within 80° of the direction directly down was removed to account for attenuation by the shadowing of the car.

7.2.8 Simulation Results and Analysis

Bit level simulations were carried out at both frequencies for each of the environments using the system parameters in Table 7.2. These parameters were chosen to fit loosely with the 5G-new-radio physical layer specification, with the length of the OTFS symbols chosen to fit within one transmission frame. Subcarrier spacing was chosen to keep intercarrier interference to a minimum. Total transmission bandwidths of 100 MHz and 20 MHz were used at 28 GHz and 3.5 GHz respectively. To remove the effects of slow fading, transmit power was chosen such that the received signal to noise ratio (SNR) was the same at each position along the route. A carrier offset was applied at the TX such that at the RX the Doppler shift of the strongest path was equal to zero.

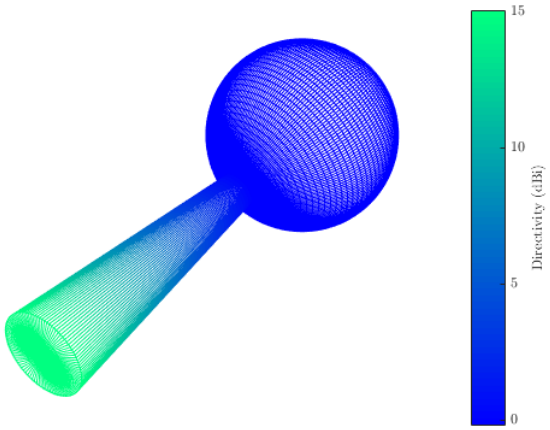


Figure 7.5: Antenna "pattern" used for 28 GHz channels.

Fig. 7.6 shows the block error rate for the motorway environment as the TX moves along the route. It can be seen that OTFS outperforms OFDM at both 3.5 GHz and 28 GHz, with a fairly constant performance gap between the two. Note that despite the constant SNR along the route, the performance of both OFDM and OTFS varies considerably, especially at 3.5 GHz. Inspection of the channels at the positions at which a high block error rate was observed found that at these locations there were multiple significant paths between TX and RX, causing deep fades in the time-frequency domain, or higher interference in the Doppler-delay domain, and therefore reducing performance. The lower variation in error in the mmWave channel could be attributed to the relative sparsity of channel, which was dominated by the line-of-sight path, due to the increased path loss of reflected paths that occurs at these frequencies and spatial filtering by the directive antenna (see Fig. 7.4).

Fig. 7.7 shows the block error rate for the town environment. OTFS again outperforms OFDM at all positions along the route, at both frequencies. Whilst with the TX and vehicles moving at lower speed, the maximum Doppler shift is lower for this environment, the built up environment provides a richer scattering environment, causing more fading/interference at both 3.5 GHz and 28 GHz.

These initial results indicate that OTFS can offer performance benefits over OFDM in automotive mobility scenarios, and that further investigation is due. Topics for future work include a performance comparison of the two waveforms using longer block codes and more detailed study of link level performance with channel estimation.

7.2.9 Conclusion

The performance of OTFS and OFDM has been compared in realistic vehicle to infrastructure 3D modelled environments, using ray-launching simulations to calculate the power, delay and Doppler shifts of propagation channels at sub-6 GHz and mmWave frequencies. Using practical rect-

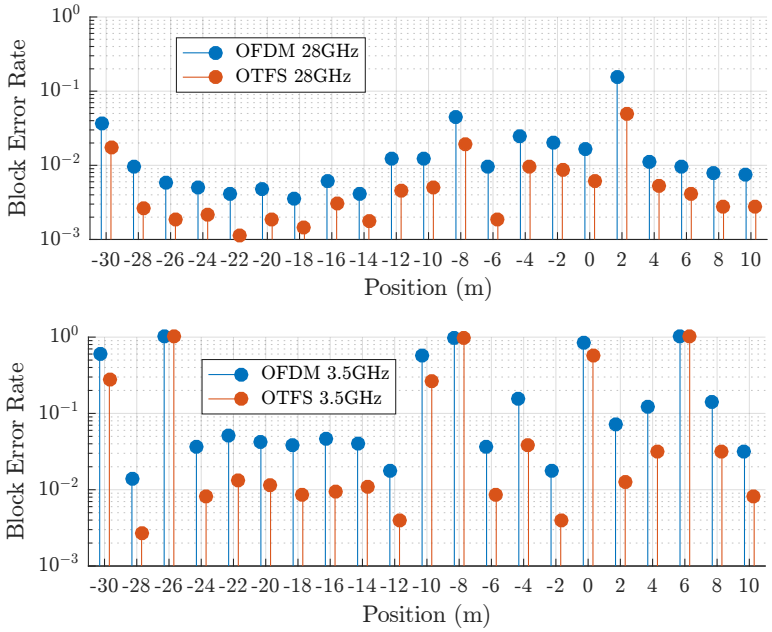


Figure 7.6: Block error rate for motorway environment, 31 m/s, at 28 GHz (top) and 3.5 GHz (bottom), with 16 QAM, 3/4 Rate, 11.5 dB SNR

angular pulse shapes and short low-density parity-check block codes, it was found that OTFS outperforms OFDM in terms of block error rate in a low mobility built-up urban environment, and high mobility motorway environment, at both 3.5 GHz and 28 GHz. The results presented show how the block error rate is consistently lower for OTFS when compared with OFDM, sometimes by an order of magnitude.

This shows that OTFS could be a useful technique in vehicular networks. It appears that the benefits are lesser at the higher frequency of 28 GHz. This may be because the channel is relatively sparser, and therefore there can be less separation of multipath components in the Doppler domain.

7.3 Investigation of a NLoS mmWave Vehicular Channel

mmWave technology can meet the high-data-rate requirement of vehicular networks. Millimetre wavelengths are a similar size to bumps on common surfaces, meaning such surfaces can be considered electrically rough. Rough surfaces diffusely scatter incident radiation. The effects of increasing the surface roughness of a primary reflecting surface, at 28 GHz, using a Kirchhoff model for rough surface scattering are demonstrated. It is shown that increasing the roughness of the reflecting surface lowers the total received power, and also increases the ratio of diffuse power to specular power scattered from the reflecting surface. Using a received bit information rate (RBIR) abstraction method, with channel data taken from a ray launching model, the effect of this scat-

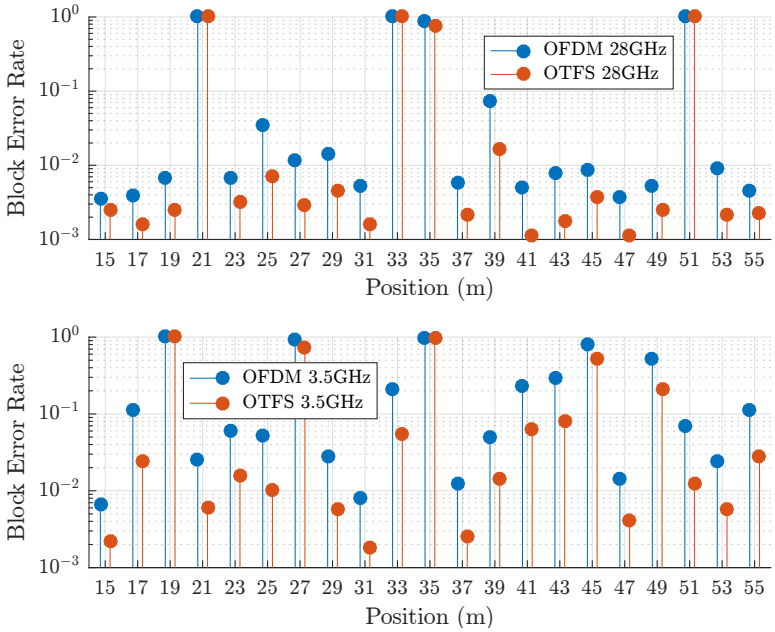


Figure 7.7: Block error rate for town environment, 13 m/s, at 28 GHz (top) and 3.5 GHz (bottom), with 16 QAM, 3/4 Rate, 11.5 dB SNR.

tering behaviour is quantified in terms of throughput. It should be noted that the RBIR simulations did not take into account the Doppler shifts calculated along the ray paths, only the channel impulse response. Roughness is shown to lead to a dramatic twofold reduction in throughput between smooth and medium roughness, and a loss of connection for rough surfaces.

7.3.1 mmWave for Vehicular Communication

Connected and autonomous vehicles will play a pivotal role in future Intelligent Transportation Systems (ITS) for smart cities. High-speed and low-latency communication links will allow municipalities to inform vehicles of safety hazards, as well as support cloud-driving solutions to reduce traffic jams and air pollution.

To meet the high capacity demands and requirements for ITS services, future vehicular communication systems will use the wide spectrum available in millimetre wave frequency bands [11]. In the UK, Ofcom has already recommended the 24.25 GHz-27.5 GHz band for ultra-dense high capacity networks, and the 57 GHz-64 GHz band is internationally designated for license-free applications, with a view to reserving the 63 GHz-64 GHz part for use in ITS applications.

7.3.2 Modelling NLoS Vehicular Channels

Autonomous vehicles will coexist with other vehicles in complex built environments, causing blockages. To communicate between vehicles and infrastructure, NLoS channels could be used.

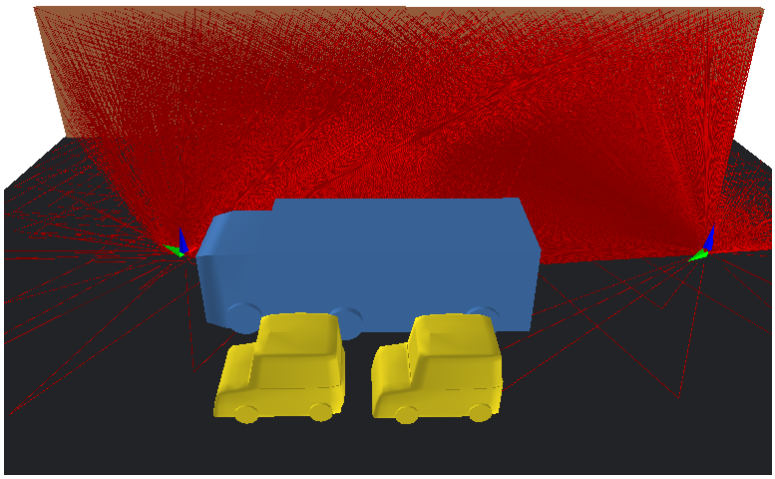


Figure 7.8: The simulated scene at a rough wall to vehicle distance of 12 m showing the road surface, rough wall and convoy of vehicles.

The feasibility of NLoS mmWave communication was demonstrated in [17]. Work based on measurements, like [17], clearly gives some insight into the mmWave channel, from which useful conclusions can be made. However it is difficult to focus exclusively on an individual propagation mechanism using traditional measurements, and in that sense, the insight they can provide is limited.

In [5][24], ray tracing was used to investigate mmWave vehicular networks, but a detailed analysis of physical propagation phenomena was not within the scope of those works. In [3], a more sophisticated ray tracing investigation of a mmWave vehicular network in a large propagation environment was conducted. However, the environment was dominated by a LoS path, and only channel metrics such as RMS delay spread were presented, making it difficult to asses the effect of mmWave specific propagation phenomena on the spatial properties of the channel.

Here, ray launching is used to investigate a mmWave NLoS channel, accounting for diffuse scattering from rough surfaces at 28 GHz. This work extends the work in [20] by including multiple reflections between two lanes of vehicles, adding motion to the scene, and by conducting a throughput analysis using the simulated channel response.

7.3.3 Simulation Scenario

Scenario

The scenario shown in Fig. 7.8 was simulated at 28 GHz. Two lanes of traffic in a convoy were simulated. The simulation included multiple reflections between vehicles, off the ground and off the reflecting wall. The wall was specified as being either smooth, medium or rough to investigate if the roughness properties of scattering surfaces affect achievable link quality.

Object	ϵ_r	μ	Velocity (m/s)
Vehicles	$1 + (1 \times 10^6)j$	1	10
Wall/Ground	$5 + 0.3j$	1	0
TX/RX	NA	NA	10

Table 7.3: Electromagnetic parameters and velocities.

The transmitter (TX) and receiver (RX) were separated by 16 m in the direction of travel. Both TX and RX were 1 m from the ground. The vehicles were considered to be metallic, with other surfaces composed of a material similar to concrete [12]. The TX, RX and vehicles were specified as moving along the road at 10 m/s. Results can easily be scaled for other velocities. Electromagnetic parameters can be found in Table 7.3 [12]. The simulation was designed to mimic vehicle mounted antennas. The vehicles attached to the TX and RX are not included in the simulation explicitly because having antennas so close to a reflecting surface in ray tracing would give spurious results. The results obtained are realistic because blockage by the missing vehicles is accounted for by nulls in the used antenna's radiation patterns when a directional pattern is applied, which will always be the case in practical mmWave systems.

The simulation was repeated at several distances from the rough wall, from 2 m to 14 m. Because diffuse scattered power "re-spreads" from a scattering element, the distance from the wall was considered to be an interesting variable affecting the significance of the diffuse scattered component. Assigning speeds to the vehicles allows for Doppler shifts to be calculated for each ray path using the method in [26].

7.3.4 System Model

The ray launching tool from [26] was used to simulate the scene using 2×10^4 rays with up to five orders of reflection and one order of diffuse scatter.

The qualitative descriptions of the wall, namely smooth, medium and rough, corresponded to the standard deviations of surface height, σ, and correlation lengths, L, as shown in Table 7.4, where σ and L have the same definition as in [4][20]. These quantities may be obtained in practice using surface scanning technology, as in [20], or using a simplified method, as in [2]. The values for the smooth and medium surfaces correspond to a material similar to concrete in a smooth or rough finish respectively [9]. The 'rough' surface was chosen to be deliberately extremely rough to illustrate a limiting case - such surfaces are unlikely to exist in practice. The parameter g is defined as [4],

$$g = \sigma^2 k^2 (\cos(\theta_1) + \cos(\theta_2))^2, \tag{7.9}$$

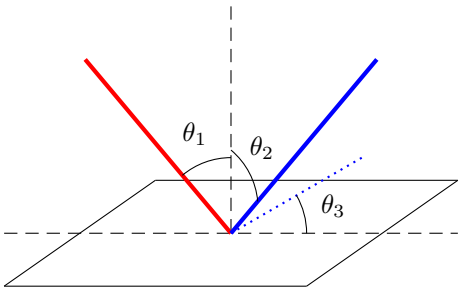


Figure 7.9: The geometry of a scattering tile showing angles θ_1, θ_2, and θ_3.

Wall	$\sigma(mm)$	$L(mm)$	$g(\sigma, k, 45°)$
Rough	7.0	22	$33.7 \gg 10$
Medium	2.4	22	$3.96 \approx 1$
Smooth	0.34	22	$0.08 \ll 1$
Ground	0.34	42	0.08

Table 7.4: Surface parameters.

where the meanings of θ_1 and θ_2 are as shown in Fig. 7.9 and k is the wavenumber. Note this is exactly the same as the definition presented in Equation 7.5 if $\theta_1 = \theta_2$. In Fig. 7.9 a square tile, to which the model is applied, is shown. The scattering model presented in this paper is applied to a rectangular tile. The size of the tile is calculated based on the size of a ray tube as it intersects a surface. The area of the tile is denoted as A. Table 7.4 shows g for an incidence angle and reflection angle of 45 degrees for each surface, showing, crudely, that the three surfaces are indeed smooth, medium and rough at the frequency of interest, according to the interpretation of the value of g as indicated in the third column of Table 7.4 [4].

The same model for losses along ray paths as Equation 7.3 of this chapter is used for the simulations in this section. Readers may refer to that section for more details.

New Method for Calculation of Tile Size

For the ray launching tool used, the area of a scattering tile was calculated based on the dimensions of the ray tube as it interacted with a surface. A ray tube could be modelled as an infinite cone with vertex angle α rads. The central axis of the cone, from the cone's vertex to its intersection point with a plane, a ray, is h units long. The angle made to a plane by the ray is β rads. See Fig. 7.10 for clarification. When calculating the area of the plane-cone intersection with $\beta < \alpha/2$, an infinite area is obtained, which gives meaningless values for the scattering coefficients. Instead, the cross sectional area of the cone, for a given value of h, is calculated. The radius of the circular cross

section, r, is given by

$$r = h \cdot \arctan\left(\frac{\alpha}{2}\right), \tag{7.10}$$

from which the cross sectional area, A_{cs} can be calculated. In [26], this area was used as the area of a scattering element. Whilst this ensures that even at glancing angles the calculations performed on the ray are representative of the field across the whole tube, it results in some parts of a surface not being completely covered by scattering elements, because no attempt is made to project the area onto the wall along the ray direction. To improve coverage in this work, the assumption is made that if h is large and α is small, the ray tube can be considered cylindrical along a short section of its central axis (i.e. $\ll h$), rather than conical. This means the area of the tile, which is the cross section of the tube projected onto the wall, can be calculated as,

$$A = A_{sc} \cdot \frac{1}{\cos(\pi - \beta)} \cdot \kappa. \tag{7.11}$$

Because ray tubes with circular cross sections overlap when projected onto a surface, too much scattered power would be predicted by the model due to double counting of some areas of the wall. To account for this the area was scaled by $\kappa = (3\sqrt{3})/(2\pi)$, which is the ratio of the area of a hexagon to the area of a circle that circumscribes it. Hexagons perfectly tessellate, so there is no double counting of area, and therefore no over-prediction of scattered power.

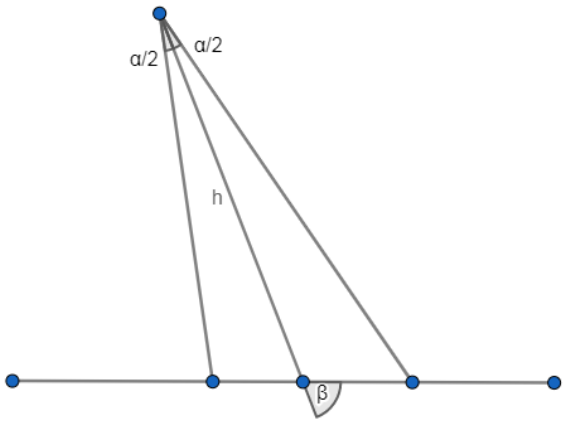


Figure 7.10: A ray tube intersecting a plane.

Antennas

Because mmWave systems will use antennas with directive radiation patterns, a simulated pattern for a circular horn antenna was used to spatially filter the channel. The bore-sight of the antenna

at both the TX and the RX was always directed at the point of specular reflection on the wall. In this way, a "dominant path" beamforming algorithm is mimicked, in which the strongest path is the reflection from the wall. The 3 dB beam-width of the antenna was 14 degrees. A transmission of vertical polarisation at the TX to vertical polarisation at the RX was considered.

Kirchhoff Model

A Kirchhoff model was used to simulate the diffuse scattering of electromagnetic waves from the rough surfaces. To show the effect of varying surface roughness, the model output for the local transverse electric to transverse electric scattering coefficient, R_{TETE}, for the three roughness cases specified in Table 7.4, is shown in Fig. 7.11. It can be observed that for the smooth wall, most energy is scattered in directions that are near the specular direction. For the medium wall, this directional behaviour is still seen, but power is scattered over a wider range of angles, and the peak scattered power is lower. For the rough wall, the scattering is almost perfectly diffuse, with a uniform spatial distribution of scattered power, this means there is no specular "peak".

These plots were generated using refractive indices of 1 and 5. The dark horizontal band on each plot corresponds to the Brewster angle for these refractive indices.

As in [14], the specular component was de-embedded from the original formulation of the Kirchhoff model [4], leaving only the diffuse component to avoid physically unreasonable results. The specular component was accounted for separately by reducing the standard Fresnel reflection matrix by the roughness correction factor,

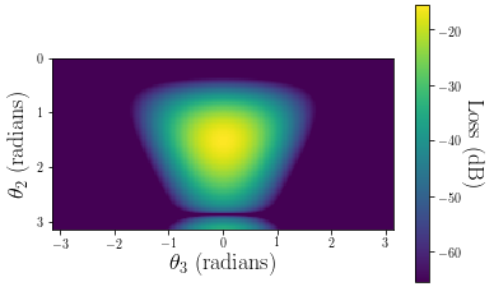
$$R_{rough} = R_{Smooth} e^{-g}, \tag{7.12}$$

where g is given in (7.9). Mathematically the Kirchhoff model for the diffuse components is expressed in section 5.4.

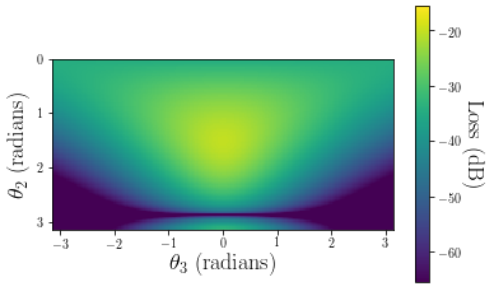
7.3.5 Results

The simulated channel was processed to calculate the total intensity (non-coherent sum) of the diffuse power shown in Fig. 7.12; the specular power, including paths with multiple reflections but no diffuse scatter, is shown in Fig. 7.13. Several clear trends are demonstrated.

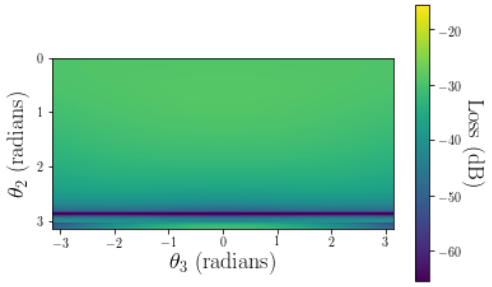
1. More specular power is received when directive antennas are used.
2. The diffuse power contained in the main beam of the antenna radiation pattern is amplified when directional antennas are used. This appears to outweigh the fact that some non specular paths are filtered by the directional pattern.
3. At all distances, the specular power is strongest for the smooth wall, less strong for the medium wall and weakest for the rough wall. The sensitivity of the specular path's power to roughness is an important result for beam-forming algorithm design.



(a) A smooth tile.



(b) A medium tile.



(c) A rough tile.

Figure 7.11: R_{TETE} for an incidence angle of 45 degrees, at a distance of 2 m, for a 10 cm by 10 cm tile relative to the power reflected in the specular direction by a perfectly smooth surface.

An interesting trend is that, unlike the specular power, diffuse power seems to remain stable regardless of the distance to the wall. This is true of both when the omnidirectional pattern is applied, and also when the directional pattern is applied. This is an unnerving result and deserves some careful analysis. In the omnidirectional case, paths scatted from the surface of the road - rather than from the wall - are dominant contributors to the diffuse power. In Fig. 7.14, the ray paths that have power contributions that are within 10 dB of the strongest ray path are shown. Higher power paths are shown in lighter blue-green, lower power paths are shown in darker blue. In Fig. 7.14a, the simulation for the smooth wall at a distance of 8 m using the omnidirectional radiation pattern is shown. It can be observed that many of the lightest coloured paths are scattering from the road surface, rather than the wall. This means that distance to the wall has little influence

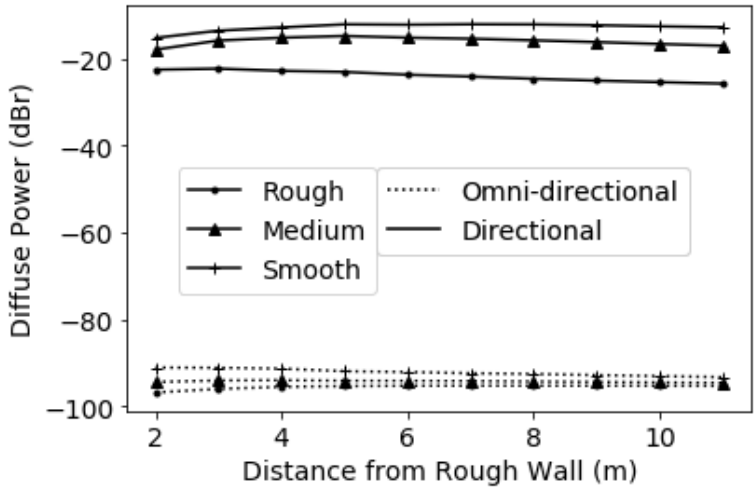


Figure 7.12: Diffuse power level at different distances to the smooth, medium or rough wall.

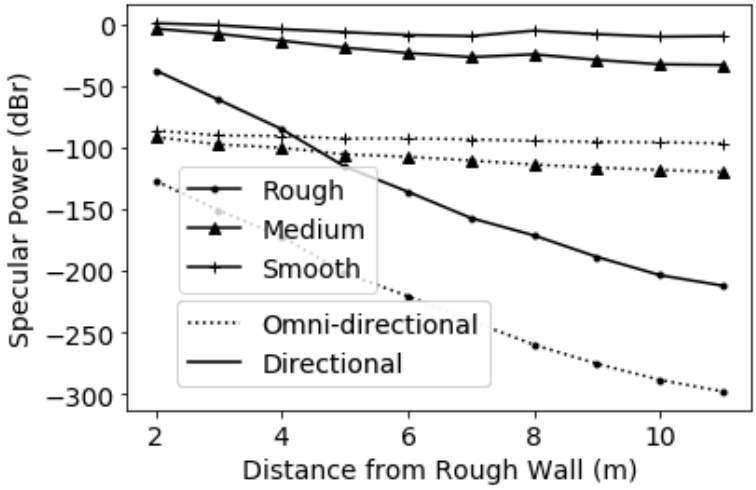


Figure 7.13: Specular power level at different distances to the smooth, medium or rough wall.

on the received diffuse power, because it has no relation to the paths from the road surface, thus explaining the behaviour shown in Fig. 7.12. In the directional case however, paths from the road surface fall on nulls in the antenna radiation pattern. This means that the mechanism causing the scattered power not to decay with distance is different. As previously stated, the TX and the RX were 16 m apart in the direction of travel. At a 2 m wall - convoy separation, free space path loss for the specular path is 86 dB, whilst at a 12 m wall convoy separation, the free space path loss is only 91 dB - a difference of only 5 dB. At greater distances to the wall, a greater area of the wall falls within the high gain region of both the antenna radiation patterns, meaning more of the wall can contribute non-specular paths that fall on a high gain region of the radiation pattern. Also, at greater separations, the geometry of the arrangement of TX, RX and scattering element mean that, for a given section of the wall, the scattering angles will be closer to specular, Figs. 7.14b and 7.14c illustrate this. Referring to Fig. 7.11, this has a profound impact on the diffuse scattered power for smooth and medium roughness surfaces, which is directional. Indeed, for the rough surface, where scattering is almost uniform, there is and observed reduction in diffuse power of just less than 5 dB moving from the closest position to the furthest, whereas for the smooth and medium roughness surfaces, diffuse power actually increases for a few steps as the convoy moves from the closest separation, and remains almost constant thereafter.

Doppler

Another metric of interest is the RMS Doppler spread. Higher Doppler spreads can cause time variation in a mobile channel, which may significantly impair systems using OFDM, or single carrier systems employing frequency domain equalisation [19]. In [26] a method was developed for calculating Doppler shifts on individual paths found by the ray launching procedure in a dynamic scene.

	Directional	Isotropic
Smooth	19	462
Medium	39	529
Rough	68	606

Table 7.5: RMS Doppler spread (Hz).

The mean RMS Doppler spread, over all distances from the wall, for the smooth, medium and rough walls is given in Table 7.5, assuming a constant vehicle speed of 10 m/s. With increasing roughness, with all other factors being constant, a greater angular spread is experienced, giving rise to a greater Doppler spread. When the directional antenna is applied, many paths are spatially filtered, and therefore the Doppler spread is reduced. However, when the directional pattern is ap-

plied, the Doppler spread for the rough wall is three times higher than for the smooth, meaning the directive antenna is less effective for reducing Doppler spread when a scattering surface is rough.

7.3.6 Throughput Analysis using Received Bit Information Rate Abstraction Method

To asses the impact of the discussed physical effects upon system performance, a throughput analysis was conducted using the received bit information rate technique in a script designed for mmWave (5G) links [1]. The RBIR script produces a throughput value based on the complex impulse response of the channel, as produced by the ray launching simulations.

A single carrier channel of 200 MHz bandwidth was investigated, which is a representative bandwidth for carrier frequencies in the range 26-28 GHz. The noise figure was specified to be 7 dB and an implementation loss of 3 dB was also included. An optimal modulation and coding scheme was selected by choosing the combination with the highest throughput for a maximum of 10 % packet error rate (PER). The MCS list for the SC 802.11.ad standard was used. Throughputs were scaled to account for the smaller bandwidth used in this paper compared with those used in 802.11.ad networks. Throughput results are shown in Table 7.6 for a transmission power of 0 dBm and in Table 7.7 for a transmission power of -15 dBm. These powers refer to the power delivered to the port of the antenna. The peak antenna gain was approximately 22 dBi, so for the 0 dB delivered power case, peak EIRP was 22 dBm. The highest achievable throughput was 428 Mb/s, which could be achieved using a 16 QAM modulation scheme with a 3/4 rate code.

7.3.7 Conclusions

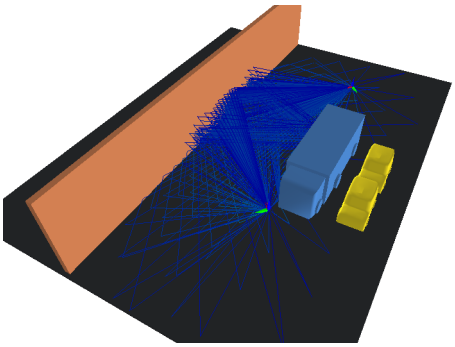
This section demonstrated the impact of rough surface scattering on the quality of the NLoS link in a vehicle-to-vehicle scenario at 28 GHz using a ray-launching analysis based on a Kirchoff scattering model. Outage was produced for a transmission power of -15 dBm when scattering from the roughest surface. Degradation to less than half the maximum achievable throughput was also experienced for the medium surface.

At a higher transmission power, of 0 dBm, the smooth and medium surfaces achieved the highest possible throughput, with the rough surface achieving a mean throughput of less than half of this for various distances of the convoy from the wall. If a convoy of vehicles was communicating via a NLoS path, and then a roadside surface changed from smooth to medium at a certain location along the road, it is likely that a significant reduction in throughput would be produced possibly prompting a search for alternative paths by a beamforming algorithm.

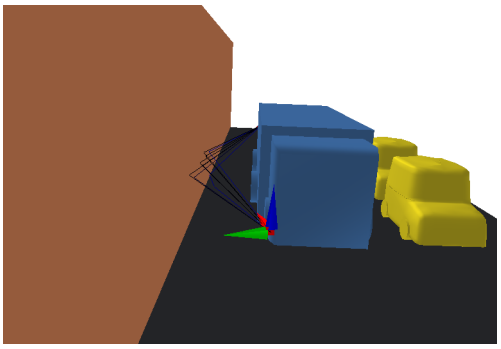
Even when transmission power was increased there was still a degradation in performance when reflecting from the roughest wall.

Distance from Wall (m)	Rough	Medium	Smooth
2	428	428	428
3	214	428	428
4	232	428	428
5	0	356	428
6	71	428	428
7	232	428	428
8	229	428	428
9	71	428	428
10	178	428	428
11	89	428	428
12	232	428	428
Mean	180	421	428

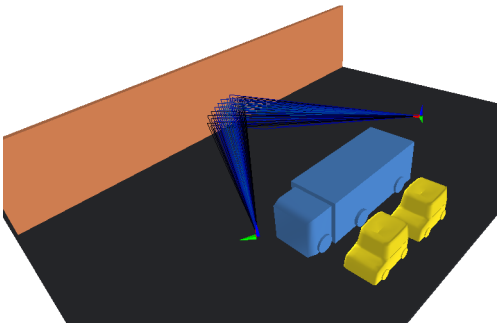
Table 7.6: Throughput (Mb/s) for rough, medium and smooth walls at 0 dBm transmission power



(a) Ray paths within 10 dB of the strongest path at 8 m separation to the rough wall using omni-directional radiation patterns.



(b) Ray paths within 10 dB of the strongest path at 2 m separation to the smooth wall using directional radiation patterns.



(c) Ray paths within 10 dB of the strongest path at 12 m separation to the smooth wall using directional radiation patterns.

Figure 7.14: Significant ray paths for three scenarios of interest.

Distance from Wall (m)	Rough	Medium	Smooth
2	36	428	428
3	0	356	428
4	0	232	428
5	0	0	356
6	0	36	266
7	0	136	285
8	0	36	232
9	0	89	346
10	0	36	232
11	0	71	89
12	0	36	214
Mean	3	132	300

Table *7.7*: Throughput (Mb/s) for rough, medium and smooth walls at -15 dBm transmission power

Bibliography

[1] N. F. Abdullah, R. Nordin, A. Doufexi, and A. R. Nix. Effect of beamforming on mmwave systems in various realistic environments. In *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, pages 1–5, June 2017.

[2] S. Ahsanuzzaman Md Tariq, C. Despins, S. Affes, and C. Nerguizian. Rough surface scattering analysis at 60 ghz in an underground mine gallery. In *2014 IEEE International Conference on Communications Workshops (ICC)*, pages 724–729, June 2014.

[3] B. Antonescu, M. T. Moayyed, and S. Basagni. mmwave channel propagation modeling for v2x communication systems. In *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pages 1–6, Oct 2017.

[4] Petr Beckmann. *The scattering of electromagnetic waves from rough surfaces / Petr Beckmann, André Spizzichino.* The Artech House radar library. Artech House, Norwood, MA, 1963.

[5] J. Choi, V. Va, N. Gonzalez-Prelcic, R. Daniels, C. R. Bhat, and R. W. Heath. Millimeter-wave vehicular communication to support massive automotive sensing. *IEEE Communications Magazine*, 54(12):160–167, December 2016.

[6] R. Hadani, S. Rakib, M. Tsatsanis, A. Monk, A. J. Goldsmith, A. F. Molisch, and R. Calderbank. Orthogonal time frequency space modulation. In *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–6, March 2017.

[7] Ronny Hadani, Shlomo Rakib, Shachar Kons, Michael Tsatsanis, Anton Monk, Christian Ibars, Jim Delfeld, Yoav Hebron, Andrea J. Goldsmith, Andreas F. Molisch, and Robert Calderbank. Orthogonal time frequency space modulation, 2018.

[8] C. Jansen, S. Priebe, C. Moller, M. Jacob, H. Dierke, M. Koch, and T. Kurner. Diffuse scattering from rough surfaces in THz communication channels. *IEEE Transactions on Terahertz Science and Technology*, 1(2):462–472, Nov 2011.

[9] J. Lin, C. R. Liu, J. Li, and X. Chen. Measurement of concrete highway rough surface parameters by an x-band scatterometer. *IEEE Transactions on Geoscience and Remote Sensing*, 42(6):1188–1196, June 2004.

[10] J. Lorca, M. Hunukumbure, and Y. Wang. On overcoming the impact of doppler spectrum in millimeter-wave v2i communications. In *2017 IEEE Globecom Workshops (GC Wkshps)*, pages 1–6, Dec 2017.

[11] Ioannis Mavromatis, Andrea Tassi, Giovanni Rigazzi, Robert J. Piechocki, and Andrew Nix. Multi-radio 5G architecture for connected and autonomous vehicles: Application and design insights. *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems*, 18(13), 3 2018.

[12] M. Olkkonen, V. Mikhnev, and E. Huuskonen-Snicker. Complex permittivity of concrete in the frequency range 0.8 to 12 ghz. In *2013 7th European Conference on Antennas and Propagation (EuCAP)*, pages 3319–3321, April 2013.

[13] S. Priebe, M. Jacob, C. Jansen, and T. Kürner. Non-specular scattering modeling for THz propagation simulations. In *Proceedings of the 5th European Conference on Antennas and Propagation (EUCAP)*, pages 1–5, April 2011.

[14] S. Priebe, M. Jacob, C. Jansen, and T. Kürner. Non-specular scattering modeling for thz propagation simulations. In *Proceedings of the 5th European Conference on Antennas and Propagation (EUCAP)*, pages 1–5, April 2011.

[15] S. Priebe, M. Jacob, and T. Kürner. Polarization investigation of rough surface scattering for THz propagation modeling. In *Proceedings of the 5th European Conference on Antennas and Propagation (EUCAP)*, pages 24–28, April 2011.

[16] S. Priebe, M. Jacob, and T. Kürner. Polarization investigation of rough surface scattering for thz propagation modeling. In *Proceedings of the 5th European Conference on Antennas and Propagation (EUCAP)*, pages 24–28, April 2011.

[17] S. Rajagopal, S. Abu-Surra, and M. Malmirchegini. Channel feasibility for outdoor non-line-of-sight mmwave mobile communication. In *2012 IEEE Vehicular Technology Conference (VTC Fall)*, pages 1–6, Sept 2012.

[18] T. S. Rappaport, Y. Xing, G. R. MacCartney, A. F. Molisch, E. Mellios, and J. Zhang. Overview of millimeter wave communications for fifth-generation (5g) wireless networks 2014;with a focus on propagation models. *IEEE Transactions on Antennas and Propagation*, 65(12):6213–6230, Dec 2017.

[19] P. Robertson and S. Kaiser. The effects of doppler spreads in ofdm(a) mobile radio systems. In *Gateway to 21st Century Communications Village. VTC 1999-Fall. IEEE VTS 50th Vehicular Technology Conference (Cat. No.99CH36324)*, volume 1, pages 329–333 vol.1, Sep. 1999.

[20] Lawrence Sayer, Alberto Loaiza Freire, Evangelos Mellios, and Andrew Nix. A kirchhoff scattering model for millimetre wavelength wireless links. In *2018 12th European Conference on Antennas and Propagation (EUCAP 2018)*, 2018.

[21] Mellios E & Nix A Sayer, L. On the feasibility of nlos vehicular communication via rough surfaces at 28 ghz. In *2020 IEEE 91st Vehicular Technology Conference*, 2020.

[22] S. Ahsanuzzaman Md Tariq, C. Despins, S. Affes, and C. Nerguizian. Rough surface scattering analysis at 60 ghz in an underground mine gallery. In *2014 IEEE International Conference on Communications Workshops (ICC)*, pages 724–729, June 2014.

[23] A. Torabi and S. A. Zekavat. Directional channel modelling for millimetre wave communications in urban areas. *IET Communications*, 12(6):656–664, 2018.

[24] Tomotaka Wada, Makoto Maeda, Minoru Okada, Katsutoshi Tsukamoto, and Shozo Komaki. Theoretical analysis of propagation characteristics in millimeter-wave intervehicle communication system. *Electronics and Communications in Japan (Part II: Electronics)*, 83(11):33–43.

[25] F. Wiffen, L. Sayer, M. Z. Bocus, A. Doufexi, and A. Nix. Comparison of otfs and ofdm in ray launched sub-6 ghz and mmwave line-of-sight mobility channels. In *2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pages 73–79, Sep. 2018.

[26] F. Wiffen, L. Sayer, A. Nix, A. Doufexi, and Z. Bocus. Comparison of OTFS and OFDM in ray launched sub-6 GHz and mmwave line-of-sight mobility channels. In *Proceedings of the 29th Conference on Personal and Indoor Mobile Radio Communications*, 2018.

Chapter 8

Conclusions & Recommendations

8.1 Aims and Scope

In contemporary popular discourse there is a suspicious attitude towards communications technology. Self parking cars, smart fridges or the latest craze in mobile gaming are seen, perhaps rightly, as peripheral, distracting and even immoral. The antisocial aspects of social media are emphasised, and the market controlled entertainment industry is vilified for promoting poor taste. However, this criticism cannot be a full account. Technology is a tool, making it easier to cause harm, but also to make genuine positive contributions to society. It is hoped by stakeholders that the effects of the technology, intended or otherwise, are positive, and where they are not, that it will be possible to address this shortcoming technically or politically. If there is good reason to be hesitant about 5G and to approach it with a healthy dose of pessimism, there are also reasons to anticipate that it will make some positive contributions to daily life. If driverless cars do improve safety, traffic flow and fuel efficiency; if spot treating crops using drones does lower harmful pesticide use; and if augmented reality does enrich our experience, then there are well founded reasons for seeking to attain technological progress. Indeed communications technology makes the working day much more effective by supplying tools with which to access rich sources of information quickly. Likewise the potential to remove the need for humans to conduct more menial tasks such as driving, means that more effective use of limited time can be made. Like many spheres of human activity, although it may be impossible to 'get right', the outcome is generally better for trying.

In response to the challenges presented by the limited and fragmented spectral resources currently available to wireless network designers, new technologies have been proposed. The main contribution of this thesis is more appropriate planning tools for these technologies. The core problem that has been addressed is that, for some of the technologies that will form a key part of 5G networks, many commonly available modelling tools are not suitable. This argument is made refer-

ring to established literature which shows that more spatial models will be required for mmWave and MIMO networks, and that traditionally popular stochastic models cannot provide enough spatial information. It was argued that a solution to this problem is the use of ray models. These inherently capture spatial information, making them suitable in an essential way. However, it was argued that ray models themselves require further development.

8.2 Ray Modelling as a System Design Tool

In chapters 2 and 3, radio propagation and propagation modelling were introduced. Chapter 2 covered the requisite theoretical background for how electromagnetic propagation could be modelled using ray optical techniques. Chapter 3 discussed some of the requirements for 5G propagation models, namely the importance of diffuse scattering, mobility and enhanced speed and accuracy.

Using a state of the art ray tracing tool, a new iterative method for antenna design was presented which made use of ray tracing results. The IoT example showed that ray tracing could verify the design of a new wearable antenna, demonstrating that modelling tools of this type could be used to improve the device design process and therefore motivating the need for work presented in subsequent chapters.

The use of the ray tracing tool greatly increased the number of links it was possible to investigate in a three storey house when compared to a manual measurement process. Although the ray tracing scenario was relatively simple, only walls and floors were included, the spatial information that was gathered is significantly more representative than a basic stochastic channel model. The example antenna design was one that was able to switch between a more or less directive radiation pattern. In the example case it was demonstrated that the ability to switch between radiation patterns was advantageous when compared to alternative designs that were only directive or only not directive. The spatial channel information was essential for quantifying the potential performance of the new design in various locations in various orientations. The proposed method captured the complex interaction of an antenna and the body with the channel which is one of its key advantages. The use of the method is recommended to those designing IoT devices or wireless sensor networks, which are both examples where a very high number of simulated channels will be useful in evaluating expected performance. This technique would be a valuable augmentation of the over the air test methods already proposed for wearable devices in which anechoic chambers or reverberation chambers are used [1]. Standard databases of test channels could be compiled allowing different designs to be compared in a controlled way, or the specialist testing of particular, more complex scenarios.

8.3 Ray Launching Tool

In Chapter 4 the design of a new ray launching tool was documented. The purpose of the new implementation was to have an experimental tool that could be easily modified to include new features, for example, new diffuse scattering models. The tool was later used for the research presented in Chapter 5 and Chapter 7. An analysis of the effects of increasing or decreasing the angular resolution of launched rays raised interesting research questions about the accuracy of phase predictions for some specific networks. It was shown that achieving phase accuracy was not necessarily possible using ray launching techniques owing to the discretisation inherent to the technique. This problem was addressed in Chapter 6. The tool was designed to produce data in a compatible format as a more mature tool developed by the author's research group [2], with the added advantage that fully 3D vector environments could be handled. The tool allowed an arbitrary number of interactions to be modelled, but was limited to one diffuse scatter as the first bounce after departing the TX or last bounce before arriving at the RX. Interactions could be reflection, refraction, diffuse scatter or diffraction.

Materials in the environment were classified with electromagnetic properties and roughness statistics using data gathered through state of the art measurement techniques or reference to literature. Various optimisation techniques were used including the use of a ray triangle intersection library and parallelism, which is essential for ray modelling large propagation environments in a timely way. Whilst it would be difficult to be better than every other academic and commercial ray tool in every respect, the developed tool had several features that would be beneficial to incorporate into the contemporary practice of ray propagation software development. The ability to be able to model overhanging features in the environment and to use vector based databases sets the developed tool apart from some other designs. The tool would allow specific scenarios to be investigated, for example communication to trains in tunnels, or to simulate specific vehicular scenarios and evaluate the effect on the reliability of a communication link. Often designers will be interested in specific channel properties or features. The ability of the tool to have different features switched on or off makes it flexible, as often a specific propagation phenomenon has a significant performance benefit.

8.4 Diffuse Scattering and Diffraction

In Chapter 5, diffraction and diffuse scattering models were introduced. The Kirchhoff model was presented as an attractive alternative to the popular effective roughness model because its scattering radiation pattern is dependant solely on physical properties of the surface rather than heuristic tuning parameters. This meant there was no requirement to 'retune' the model against measurements for every new material or frequency of interest. A new method of finding multiply diffrac-

ted paths in a ray launcher was presented and both scattering and diffraction models were verified against measurements. A main conclusion of the work presented in this Chapter is that it is very important to reduce a specularly reflected component to account for roughness, and that a few privileged directions around the specular direction will contribute significant diffuse scattered components, however at angles far from the specular direction, diffuse scattered components will be insignificant. This conclusion may help reduce the complexity of ray tracing simulations that include diffuse scattering. Future work might include a sensitivity analysis of the effects of reducing the number of scattering locations based on various geometrical criteria and assessing the effect of both prediction accuracy and execution speed. Another important result is that diffuse scattering behaviour is much more sensitive to roughness over a plausible range than other parameters like permittivity. This indicates that some effort must be made to classify the roughness of surfaces in environment models. This presents a serious modelling challenge as roughness statistics are, in general, hard to obtain, especially for large scale environment models. This problem is mitigated somewhat by the smaller scale of mmWave networks, but remains a challenge. A model was presented that is frequency scaleable, requiring only roughness data. The use of this model would remove some 'guesswork' inherent to other models which must be retrained using measurement data.

8.5 Ray Tracing for MIMO Simulation

In Chapter 6, a ray tracing tool specifically designed for MIMO simulation was presented. The Chapter addresses some solutions to the problems identified in Chapter 4, namely that any kind of spatial discretisation causes errors in path length, and therefore errors in phase. The premise of the work was that ray tracing using the image method, rather than ray launching, avoids problems due to errors in path length. However, it was shown that ray tracing computational complexity increases exponentially with the number of reflections, and that this can lead to very long run times. The main work presented in the Chapter was a novel optimisation technique to mitigate the longer run times. The technique solved the 3D intra-mesh visibility problem using a non-linear problem solving library MIDACO. It was shown the technique was superior to several alternatives in common use, such as the backface culling method, the circumcentre only method or the random point method. An extra attempt to optimise ray tracing for MIMO networks, making use of the relative proximity of array elements was made. Unfortunately, for the number of reflections allowed, these techniques were unsuccessful. Finally an evaluation of the plane wave array assumption was made. It was shown that for large MIMO arrays, the assumption is not accurate over a distance comparable to a plausible size for a MIMO antenna array.

Calculating the mutual visibility of primitives in an environment is a way of radically increasing the speed of ray tracing simulations. In current literature, this is often achieved in a nonrobust way.

The method presented in this thesis is superior to previous techniques as it does not require the use of ray launching, nor does it overestimate mutual visibility whilst converging faster to the correct minimal visibility solution faster than a random method of visibility sampling.

8.6 Ray launching Investigation of Mobility Channels

In chapter 7, the ray launching tool developed described chapter 5 was extended to include calculation of Doppler shifts along ray paths. This allowed vehicular networks including a high degree of mobility to be investigated. In the first instance, the extracted Doppler-delay power spectrum from simulations of a motorway scenario and of a city scenario were used to do a system level evaluation of the OTFS waveform, comparing results to OFDM in two mobility scenarios. In the second case the tool was used to evaluate the plausibility of NLoS vehicular communication at mmWave. It was shown that due to the sparsity of the mmWave channel, Jakes' spectrum is a poor model of a realistic Doppler spectrum. It was also shown that NLoS vehicular communication is contingent upon the visibility of relatively smooth surfaces to act as single reflectors since rough ones attenuate the signal too far. The tool was updated to seamlessly update parts of the scene that were mobile. This allowed a scene to be set up and 'let go', rather than meticulous modelling of initial and subsequent snapshots of the scene.

8.7 Final Remarks

To answer important research questions appropriate tools and models are needed. This thesis has taken an approach in which the particular research questions posed by the development of 5G technologies have influenced the design of the tools used to evaluate them. The use of these tools then helps to identify key features of the channel, showing some mechanisms are important and some less so. This in turn gives further insight into how better tools might be designed to tackle further research questions. This thesis has identified some interesting channel properties, such as realistic Doppler spectra, and the feasibility of NLOS mmWave communication in the presence of varying roughness. In doing so some important properties of the tools themselves have been explored, for example, one possible conclusion is that for resource constrained simulations, not modelling non-specular scattering at all might be prudent. On the other hand reducing the amount of power scattered in the specular direction to account for roughness is of paramount importance. The method presented in this thesis, to implement models of specific propagation phenomena, then test these against measurements and then to combine these verified models to simulate more complex scenarios, is one that is recommended for future modelling efforts as a practical balance between usefulness and rigour.

Bibliography

[1] Kelly Hill. Testing the internet of things: making the iot work. `https://www.2j-antennae.com/files/1479994838.pdf`.

[2] T. A. Thomas, F. W. Vook, E. Mellios, G. S. Hilton, A. R. Nix, and E. Visotsky. 3d extension of the 3gpp/itu channel model. In *2013 IEEE 77th Vehicular Technology Conference (VTC Spring)*, pages 1–5, June 2013.